Macro Data: Some Sources and Tools

Quantitative Macroeconomics [Econ 5725]

Raül Santaeulàlia-Llopis

Washington University in St. Louis

Spring 2016

Defining Trends and Cycles

- Many time series of data in which we are interested in grow.
- Do they grow at a stationary rate or not?
- Alternative answers to that question are in turn associated to alternative definitions of trends and cycles.
- The definition of trend and cycle that we take has implications for the way we stationarize models, and also for the model equilibrium allocations and quantitative results.
- Usually, all time-series we work with are seasonally-adjusted (X-12-ARIMA, or monthly dummies).

Deterministic Trend, DT

The DT specification of a time series is[1]

$$y_t = \chi_y + \omega_Y \; t + \widehat{y}_t \tag{1}$$

where y_t is in logs, the intercept χ_y is the estimated initial value (if t starts at 0 and $\widehat{y}_0 = 0$), the slope ω_y is the trend, and $\widehat{y}_t$, are the log-deviations from trend that we can uncover by estimating (1).

[1] That is, $Y_t = Y_0 \, (1 + \omega_Y)^t \; e^{\widehat{y}_t}$, where $Y_0 = e^{\chi_y}$, and $y_t = log(Y_t)$.

Remarks,

- In the case of a linear trend,
  - $\omega_Y = \lambda_Y$, the stationary growth rate.
  - $\lambda_{Y_{t,t-1}} = \lambda_Y + (\widehat{y}_t - \widehat{y}_{t-1})$

- Trend may be non-linear,

$$y_t = \chi_y + f(t;\Omega) + \widehat{y}_t \text{ say } y_t = \chi_y + \omega_{Y,1}\, t + \omega_{Y,2}\, t^2 + \widehat{y}_t$$

- There may exist trend breaks: productivity slowdown in 1974 (the Bai-Perron test, (Bai and Perron (1998)), computes multiple structural breaks).

- Some properties on the persistence and variance of the process determine whether the movements around the trend, $\widehat{y}_t$, are **weakly covariance stationary**.[2]

- If so, classical estimation and inference procedures apply.

- We can represent $\widehat{y}_t$ as a stationary process, perhaps AR(n),

$$\widehat{y}_t = \gamma \, \widehat{y}_{t-1} + u_{\widehat{y}_t}, \quad u_{\widehat{y}_t} \sim N\left(0, \sigma^2_{u_{\widehat{y}}}\right) \tag{2}$$

- AR specifications may have more lags. There are some criteria to choose the number of lags (Akaike (AIC), Schwarz's Bayesian (SBIC), Hannan and Quinn (HQIC), etc.). Also, if $u_{\widehat{y}_t}$ follows a moving average process, then $\widehat{y}_t$ follows an ARMA process.

[2] That is, constant variance, constant expectations, and covariance(x_t, x_{t+h}) depends on h, but not on t.

Stochastic Trend, ST

A model with a stochastic trend is an integrated process of some order. The simplest is an $I(1)$ process like a random walk,[3]

$$y_t = y_{t-1} + \lambda_Y + \widetilde{y}_t \tag{3}$$

where y_t is in logs, λ_Y is average growth rate, and $\widetilde{y}_t$ are deviations from mean growth rate. That is, y_t follows a random walk with drift λ_Y.

Here, detrending implies recovering $\widetilde{y}_t$, that defines fluctuations around average growth rate. To recover $\widetilde{y}_t$, we have to difference the data once. $\widetilde{y}_t$ is now stationary and we can estimate it,

$$\widetilde{y}_t = \widetilde{y}_{t-1} + u_{\widetilde{y}_t}, \quad u_{\widetilde{y}_t} \sim N(0, \sigma^2_{u_{\widetilde{y}}}) \tag{4}$$

and

[3] That is, $Y_t = Y_{t-1} e^{\lambda_Y + \widetilde{y}_t}$

Remarks,

- $\widehat{y}_t$ and $\widetilde{y}_t$ are related by: $\widehat{y}_t = \sum_{\tau=0}^{t} \widetilde{y}_\tau$ and, hence, $\widehat{y}_t - \widehat{y}_{t-1} = \widetilde{y}_t$. That is, we can always transform one into the other. To see this, substitute (3) backwards to $t = 0$.

- However, the artificial series of y_t that results from the two stochastic processes (2) and (4) is going to be—potentially—very different because innovations have permanent effects on the levels of y_t when we deal with stochastic trends, while they have only transitory effects on those levels when we deal with deterministic trends.

- While a deterministic trend is constant for the whole series, a stochastic trend can shift signs—positively or negatively.

- Which specification should we choose, **DT or ST?**.

- The equivalent of this question is: Is there a unit root, or not? Note that with $\widehat{y}_t = \gamma \widehat{y}_t + u_t$, if $\gamma = 1$, then $\widehat{y}_t - \widehat{y}_{t-1} = u_{\widehat{y}_t}$. Further, recall that $\widehat{y}_t = \sum_{\tau=0}^{t} \widetilde{y}_\tau$ and, hence, if $\gamma = 1$ then $u_{\widehat{y}_t} = u_{\widetilde{y}_t}$. That is, if we estimate the DT process and find that $\gamma = 1$, then we say that $\widehat{y}_t$ follows a unit root and we are back to the ST process.

- Problem: it is difficult to test. Unit root tests have low power: they can easily reject the null of stochastic trend when it should have been accepted (specially when there is a trend break in the data, see Perron (1990)). Some of these unit root tests are the Phillip-Perron test and Dickey-Fuller test.

For example, what about the stationarity of hours per capita?:

- Francis and Ramey (2001) tests of a unit root in per capita hours are not rejected at conventional significance levels. However, bear in mind that the nonrejection of a null hypothesis does not mean the alternative hypothesis, in this case level stationarity, is rejected.

- Indeed, the Kwiatkowski, Phillips, Schmidt and Shin (1992) test of the null of level stationarity against the alternative of a unit root. They find that stationarity is not rejected for at conventional significance levels either.

- Similarly neither level nor difference stationarity of per capita hours can be rejected by classical statistical criteria, see Christiano, et. al. (2004). Christiano, et. al. (2004) use Bayesian methods to argue that the preponderance of the evidence points toward level stationarity as the most plausible assumption.

In all, Unit Roots: Do We Know, Do We Care? No. May be not (Christiano and Eichenbaum (1988)).

How much should an innovation to real output affect future real output into the infinite horizon?

If the answer is zero, then real output is trend stationary. If the answer is not zero, then real GNP is difference stationary. That is, the difference is completely summarized by the answer to the question.

Christiano and Eichenbaum (1988): "The competing hypotheses have no other testable differences. Once we pose the question in this way, it seems clear that economists ought to be extremely skeptical of any argument that purports to support one view or the other."

Hodrick and Prescott Filter

- Any series y_t is composed by trend and cycle, $y_t = \tau_t + c_t$
- Given a positive value of λ, there is a unique τ_t that minimizes,

$$\min_{\tau_t} \sum_{t=1}^{T} (y_t - \tau_t)^2 + \lambda \sum_{t=2}^{T-1} \left[(\tau_{t+1} - \tau_t) - (\tau_t - \tau_{t-1})\right]^2$$

  - If $\lambda = 0$, then $\tau_t = y_t$; that is, the trend τ_t is defined as the series y_t itself, and the cyclical component is zero.
  - If $\lambda = \infty$ then the trend component τ_t is a linear trend. Hence, the cycle component is (if y_t is in logs) the log-deviation from trend.

- The first term, $y_t - \tau_t$, penalizes the cyclical component in a particular manner (more penalty to larger deviations).

- The second term penalizes changes in the growth rate of the trend, τ_t.

The standard advise is to set $\lambda = 1600$ when dealing with quarterly data, $\lambda = 14400$ for monthly data and $\lambda = 100$ for annual data.

What these values of λ do is define business cycles as those cycles shorter than 8 years. Fluctuations that are not long run and that last longer than 8 years define the medium run.

Ravn and Uhlig (2002) study how the Hodrick-Prescott filter should be adjusted when changing the frequency of observations.

- They show that the filter parameter should be adjusted by multiplying it with the fourth power of the observation frequency ratios.
- This yields an HP parameter value of 6.25 for annual data given a value of 1600 for quarterly data.

Baxter and King Filter (1999)

- While an ideal high-pass filter (as the HP filter) removes low frequencies from the data, an ideal band-pass filter removes both low and high frequencies.
- The BK is a band-pass filter. It decomposes data into three components: the trend (low frequencies), the cycle (intermediate frequencies) and irregular (high frequencies).
- Intermediate frequencies represent the cyclical component or the business cycle.

- The BK filter removes all fluctuations at both low and high frequencies. It is designed to pass through components of time series with fluctuations between 6 and 32 quarters while removing higher and lower frequencies. The remaining part is our measure of the business cycle.

- When applied to to quarterly data, the band-pass filter proposed by Baxter and King takes the form of a 24-quarter moving average,

$$y_t^f = \sum_{h=-12}^{12} a_h \, y_{t_h} = a(L)y_t$$

  where L is the lag operator.

- Note that when using the BK filter 12 quarters are sacrificed at the beginning and the end of the time series and this may limit its analysis for contemporaneous data.

Christiano and Fitzgerald Filter (2003)

- Christiano and Fitzgerald (2003) build a filter using two new ingredients :
    - Accounting for the assumed spectral density of the original data, and
    - Dropping the stationarity and symmetry conditions on the filter coefficients.

Comparing Definitions of the Cycle

- There are other methods to decompose trend and cycles (Unobserved Components models, the Beveridge-Nelson decomposition...)
- The methods we have described (or others) may decompose trend and cycle very differently.
- **Exercise:** compare the cycle defined by the H-P filter, Uhlig's modification of it, the Baxter and King Filter, and the Fitzeral and Scott Filter and discuss the resemblance to the NBER chronology. Document the years peak-to-trough and trough-to-peak associated to each filter for each NBER-recession.
- Greenwood and Yorokoglu (1997) discuss a trend break in 1974. Some people argue that if this trend break is not accounted for, it may impose large trend fluctuations (see Canova et al. (2010)). **Exercise:** Redo the above exercise but for two subsamples, pre-1973.IV and post-1974.I.

Table: Cyclical Behavior of U.S. Data 1954.I–2004.IV

	U.S. Data		
	σ_x^2	$\rho(y, x)$	$\rho(x_{t-1}, x_t)$
y	**2.53**	1.00	.85
eh	**2.43**	**.88**	.89
c	1.56	.87	.86
i	52.27	.91	.80
r	.01	.74	.78
w	.58	.08	.70
z^0, z^1	.72	.74	.70
ls	.46	-.24	.78

Notes: Logged (except for r) and HP-filtered data.
Source: Ríos-Rull and Santaeulàlia-Llopis (2010).

Table: Phase-Shift of the U.S. Data 1954.I–2004.IV

	Cross-correlation of y_t with										
	x_{t-5}	x_{t-4}	x_{t-3}	x_{t-2}	x_{t-1}	x_t	x_{t+1}	x_{t+2}	x_{t+3}	x_{t+4}	x_{t+5}
	U.S. Data 1954.I–2004.IV										
y	-.04	.14	.37	.63	.85	1.00	.85	.63	.37	.14	-.04
eh	-.22	-.06	.15	.40	.67	.88	.91	.81	.63	.41	.21
c	.14	.33	.51	.70	.84	.87	.71	.50	.26	.03	-.14
i	.05	.20	.39	.60	.79	.91	.75	.51	.24	-.01	-.22
r	.16	.31	.48	.63	.73	.74	.47	.17	-.09	-.29	-.40
w	.18	.19	.19	.17	.10	.08	-.06	-.11	-.15	-.12	-.12
s^0	.25	.39	.54	.68	.73	.74	.39	.08	-.18	-.33	-.43
ls	-.20	-.26	-.32	-.34	-.33	-.24	.03	.25	.40	.47	.44

Notes: Logged (except for r) and HP-filtered data.
Source: Ríos-Rull and Santaeulàlia-Llopis (2010).

Labor Market Puzzles

1. Dunlop-Tarshis phenomenon: $\rho(w, y) \approx 0$. See Hansen and Wright (1992) and Abraham and Haltingwanger (1995)
2. Labor productivity puzzle: low $\rho(lp, y)$. See Gali and Van Reens (2010) and McGrattan and Prescott (2012).
3. Labor share puzzle: negative $\rho(ls, y)$. See Boldrin and Horvath (1995), Gomme and Greenwood (1995) and Rios-Rull and Santaeulalia-Llopis (2010).
4. The hours-productivity puzzle: Negative $\rho(eh, w)$ and $\rho(eh, lp)$. See Christiano and Eichenbaum (1992).
5. (Un)employment volatility puzzle for DMP models: See Shimer 2005 and Pissarides (2009).

We will see (you will compute it) how business cycle models can perform in these dimensions.

Some Nonstationary Statistics

- Some data relationships are non-stationary. One way to uncover these non-stationarities is by computing moving averages of adequate intervals:
    - Rolling windows of the variances.
    - Rolling windows of the correlation between two series.
    - Rolling windows of the phase-shift two series.

- If we are interested in business cycles a common window is 10 years.

Using National Income and Products Accounts and BLS

Raw data series can be retrieved from the Bureau of Economic Analysis (BEA; www.bea.gov) and the Bureau of Labor Statistics (BLS; www.bls.gov).

National Income and Product Accounts (NIPA-BEA)

- Table 1.7.5: Gross National Product (GNP), Consumption of Fixed Capital (DEP), Statistical Discrepancy (SDis)
- Table 1.12: Compensation of Employees (CE), Proprietor's Income (PI), Rental Income (RI), Corporate Profits (CP), Net Interests (NI), Taxes on Production (Tax), Subsidies (Sub), Business Current Transfer Payments (BCTP), Current Surplus of Government Enterprises (GE)
- Table 5.7.5: Private Inventories (Inv)

Fixed Asset Tables (FAT-BEA)

- Tables 1.1 and 1.2: Private Fixed Assets (KP), Government Fixed Assets (KG), Consumer Durable Goods (KD)
- Tables 1.3: Depreciation of Private Fixed Assets (DepKP), Depreciation of Government Fixed Assets (DepKG), Depreciation of Consumer Durable Goods (DepKD)

Current Establishment Survey (CES-BLS)

- Employment (E): Series ID CES0000000081
- Hours per worker, Average Weekly Hours (h): Series ID CES0500000082, Series ID EEU00500005

The labor input (aggregate hours) is computed as the product,

$$H_t = E_t \; h_t$$

Alternatively **BLS** provides an index for aggregate hours that combines information from CES and CPS.

Constructing the Labor Share

- The labor share of income is defined as one minus capital income divided by output.

- Several sources of income, mainly proprietor's income, cannot be unambiguously allocated to labor or capital income.

- Remark: Labor Share is ratio and we use nominal variables to construct it in order to prevent distortions generated by possibly different deflators.

To deal with this we can assume that the proportion of ambiguous capital income to ambiguous income is the same as the proportion of unambiguous capital income to unambiguous income:

1. Unambiguous Capital Income (UCI) = RI + CP + NI + GE
2. Unambiguous Income (UI) = UCI + DEP + CE
3. Proportion of Unambiguous Capital Income to Unambiguous Income: $\theta_P = \frac{\text{UCI+DEP}}{\text{UI}}$
   Then we can use θ_P to compute the amount of ambiguous capital income in ambiguous income,
4. Ambiguous Income (AI) = PI + Tax - Sub + BCTP + SDis
5. Ambiguous Capital Income (ACI) = $\theta_P \times$ AI

- Then, capital income (service flows of private fixed capital), Y_{KP}, is computed as the sum of unambiguous capital income, depreciation, and ambiguous capital income, that is,

$$Y_{KP} = \text{UCI} + \text{DEP} + \text{ACI}, \tag{5}$$

which we use to construct our baseline labor share.

- Baseline labor share in Ríos-Rull and Santaeulàlia-Llopis (2010):

$$\text{LS} = 1 - \frac{Y_{KP}}{Y} \tag{6}$$

Extending the Labor Share to Durables and Government

- Assuming that the return on capital, i, is the same for fixed private capital, consumer durables, and government stock.

  (a) *KP* Fixed private capital includes producers' durables, producers' structures, and residential structures. *Note that it does not include consumers' durables, only producer's.*

  (b) *KG* Government capital stock includes both equipment and structures at both the federal level and the sate and local government level.

  (c) *KC* Consumer durables.

  1. First, we determine the return on capital, i, using only private fixed capital and solving the following equation that relates capital income to capital stock:[4]

  $$Y_{KP} = i \times (\mathrm{KP} + \mathrm{Inv}) + \mathrm{DEP} \tag{7}$$

  Note that here we are only using private fixed capital KP (+ inventories because they are a factor of production). Cooley and Prescott (1995) also add land capital from the Flow of Funds, although they suggest it is quite volatile in terms of GDP (the land data is not longer published, quality issues).

  2. The depreciation rates of consumer durables and government stock are computed as

  $$\delta_D = \frac{\mathrm{DepKD}}{\mathrm{KD}} \qquad \delta_G = \frac{\mathrm{DepKG}}{\mathrm{KG}} \tag{8}$$

[4] We transform the annual capital stock and depreciation series provided by FAT-BEA to a quarterly series by interpolation.

- This way, the flow of services from consumer durable goods and government capital can be derived as

$$Y_{KD} = (i + \delta_D) \times \text{KD} \qquad\qquad Y_{KG} = (i + \delta_G) \times \text{KG} \tag{9}$$

- Finally, the labor share with durables that extends measured output and capital income with flow services from consumer durables is

$$\text{LS w/ Durables} = 1 - \frac{Y_{KP} + Y_{KD}}{Y + Y_{KD}} \tag{10}$$

and the labor share with durables and government that also includes flow services of government stock is

$$\text{LS w/ Durables \& Gov.} = 1 - \frac{Y_{KP} + Y_{KD} + Y_{KG}}{Y + Y_{KD} + Y_{KG}} \tag{11}$$

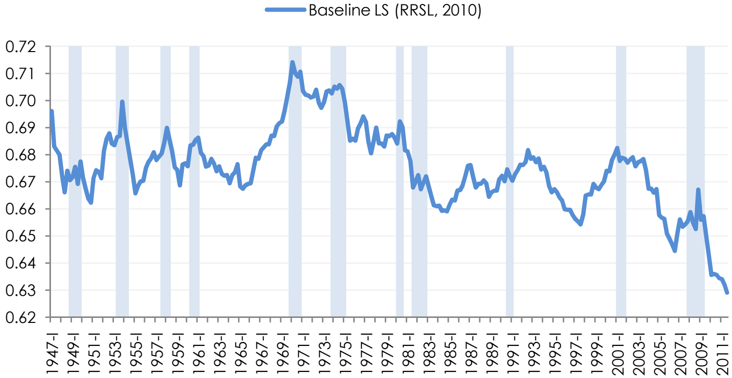


Figure: Baseline Labor Share 1947.I-2011.III

Source: Ríos-Rull and Santaeulàlia-Llopis (2010), Updated series

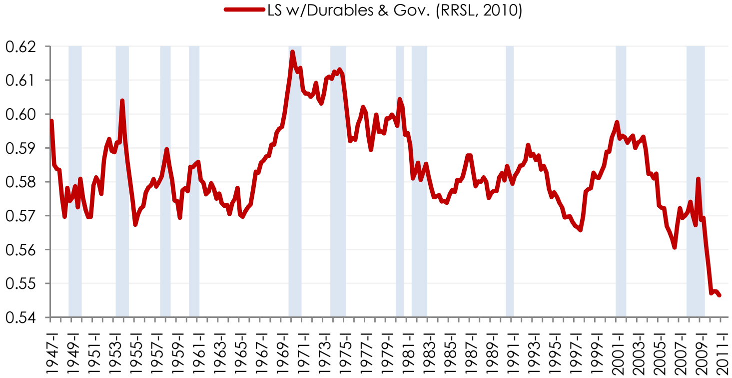


Figure: Labor Share with Durables and Government 1947.I-2010.IV

Source: Ríos-Rull and Santaeulàlia-Llopis (2010), Updated series

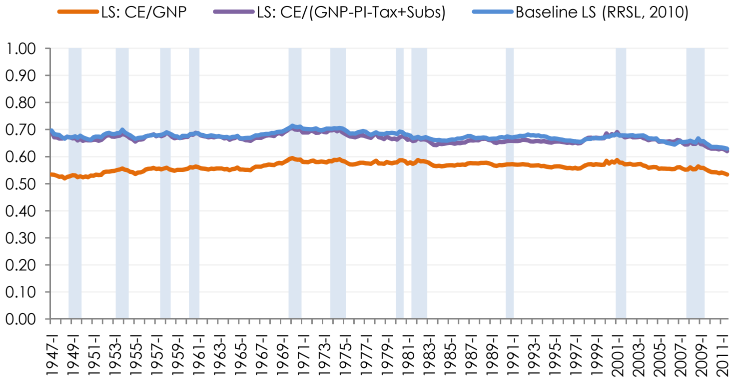


Figure: Comparison of Labor Share Definitions (I)

Source: Ríos-Rull and Santaeulàlia-Llopis (2010), Updated series

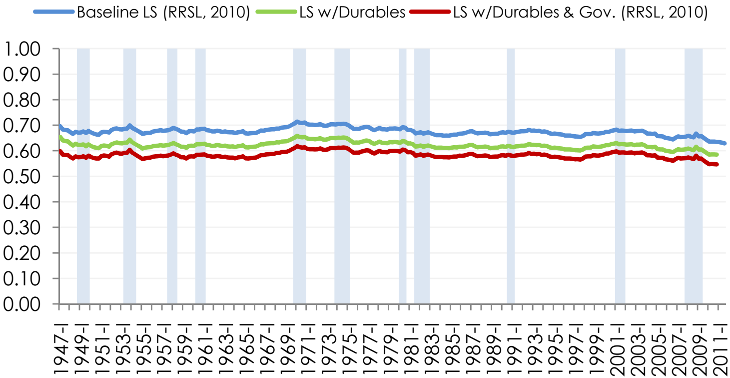


Figure: Comparison of Labor Share Definitions (II)

Source: Ríos-Rull and Santaeulàlia-Llopis (2010), Updated series

Wages

- One way to compute a time-series for wages is recovering it as a residual of data objects that we know.

- Labor Income divided by hours,

$$w_t = \frac{\text{Labor Income}_t}{N_t} = \frac{Y_t - Y_{KP_t}}{N_t}$$

or, equivalently,

- Labor Share divided times labor productivity,

$$w_t = \text{LS}_t \, \frac{Y_t}{N_t}$$

Note from above, that alternative definitions of labor income imply alternative definitions of total income and labor share. We can proceed analogously to recover series for the interest rate using series of K_t, Y_t, and $1 - \text{LS}_t$.

Constructing the Relative Price of Quality-Adjusted Investment

- Greenwood, Hercowitz, and Krusell (1997) and Greenwood, Hercowitz, and Krusell (2000) propose investment-specific technical change (ISTC) is a source of growth and business cycle fluctuations.

- The one-sector version of their model implies the identification that investment-specific technical change is the relative price of quality-adjusted investment in terms of consumption, see Fisher (2006).

- A common concern is that FAT-NIPA provides a trend for the index of equipment investment that is not adjusted by quality, see Gordon (1990) and Cummins and Violante (2002).

- Here we follow Cummins and Violante (2002) and Fisher (2006) to construct a series of quality-adjusted total capital.
- Then we use it to recover a productivity residual modified by quality.
- See appendix in Ríos-Rull et al. (2012) for details.

Additional Data Series that We Need

National Income and Product Accounts (NIPA-BEA)

- Table 1.1.5: Consumption of Durable Goods (CD_t), Change in Inventories (ChInv_t)
- Table 2.3.3 and 2.3.5: Quantity Index (QCONS_t^i) and Nominal (CONS_t^i) Nondurables Consumption (excluding Energy) and Services (excluding Housing)[5]
- Table 3.9.5: Government Investment in Equipment (GovIEQ_t), Government Investment in Structures (GovIST_t)
- Table 5.3.5: Private Fixed Investment in Equipment (PrivIEQ_t), Private Fixed Investment in Structures (PrivIST_t)

[5] Goods i correspond to nondurables consumption in food, clothing and shoes, and others, and services in households operations, transportation, medical care, recreation and others.

Fixed Asset Tables (FAT-BEA)

- Tables 5.3.4: Official Price Index for Investment in Equipment ($OPIEQ_t$)

Cummins and Violante (2002), 1947-2000

- Annual Quality-Adjusted Price Index for Investment in Equipment ($\text{QAPIEQ}^{CV}_{year}$)
- Annual Quality-Adjusted Depreciation Rates for Total Capital (δ^{CV}_{year})

The construction of the relative price of quality-adjusted investment, P_t , requires several steps, see appendix in Ríos-Rull et al. (2012). Roughly, we need to compute the following items:

1. Quarterly Price Index for Investment in Structures, PCONS_t.
2. Quarterly Quality-Adjusted Price Index for Investment in Equipment, QAPIEQ_t.
3. Quarterly Quality-Adjusted Price Index for Total Investment, QAPI_t.
4. Quarterly Relative Price of Quality-Adjusted Investment, P_t,

$$P_t = \frac{\text{QAPI}_t}{\text{PCONS}_t}$$

5. Its inverse,

$$V_t = \frac{1}{P_t}$$

is **investment-specific technical change,** V_t.

Compute Quality-Adjusted Capital Series

1. Quarterly Quality-Adjusted Investment, X_t,

$$X_t = \frac{\mathsf{InvEQ}_t + \mathsf{InvST}_t}{\mathsf{QAPI}_t}$$

2. Quarterly Quality-Adjusted Depreciation Rates, δ_t.

3. Quarterly Quality-Adjusted Capital Stock, K_t. With an initial capital stock in efficiency units, K_0, we can construct the series of capital in efficiency units recursively using the perpetual inventory

$$K_{t+1} = (1 - \delta)\, K_t + X_t$$

Productivity Residual

- We globally approximate the production function as

$$Y_t = S_t \, f(K_t, \; N_t) \tag{12}$$

- We recover the productivity residual, S_t, by totally differentiating (12). Then we can use time series of real output y_t, real capital k_t, and labor input n_t to get,

$$\lambda_{s_t} = \lambda_{y_t} - (1 - \zeta) \, \lambda_{k_t} - \zeta \, \lambda_{n_t} \tag{13}$$

where λ_{x_t} is the growth rate of variable x_t, and ζ is a relative input share parameter that, under the assumption of competitive markets, we chose to match the long-run average of the labor share of income.

- Given an initial value for S_0, we can recover the whole series.

- But s_t^0 has trend, and we want a trendless object. Consider now a detrending procedure that uses the following linear regression:

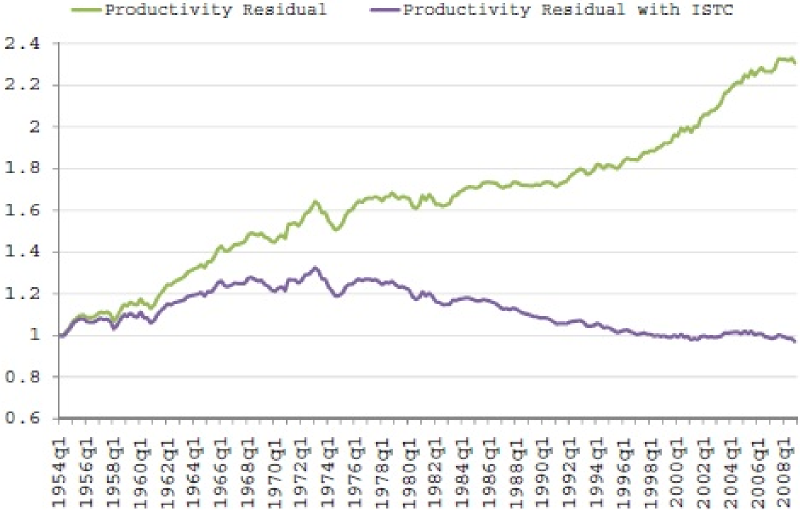
$$\ln X_t = \chi_x + \omega_x t + \widehat{x}_t, \tag{14}$$

  where X_t is any economic variable, and where χ_x and ω_x are the mean and trend parameters and $\widehat{x}_t$ are the residuals.

- To detrend the productivity residual, we can either apply (14) directly to S_t, or equivalently[6], apply (14) to the series of output, capital, and labor input to obtain the following detrended productivity residual:

$$\widehat{s}_t = \widehat{y}_t - \zeta\, \widehat{k}_t - (1-\zeta)\, \widehat{n}_t. \tag{15}$$

[6] The equivalence relies on the fact that output per capita and capital per capita have similar long-run growth in the data.


Productivity Residual
Productivity Residual with ISTC
2.4
2.2
2
1.8
1.6
1.4
1.2
1
0.8
0.6
1954q1
1956q1
1958q1
1960q1
1962q1
1964q1
1966q1
1968q1
1970q1
1972q1
1974q1
1976q1
1978q1
1980q1
1982q1
1984q1
1986q1
1988q1
1990q1
1992q1
1994q1
1996q1
1998q1
2000q1
2002q1
2004q1
2006q1
2008q1

VARs

- Introduced by Sims in the 1970s and 1980s.

- What questions can 'in principle' be adressed by VAR?
    - How does an economy respond to a particular shock? IRFs may help to answer this.
    - How much an economy responds to a particular shock? Through FEDVs or other decompositions.

There is an unsettled debate on how 'useful' or not VARs are in doing any of the above (see the debate between Christiano et al. (2006)) and with Chari et al. (2008)).

Currently, state-of-the-art estimation techniques use data IRFs as matching moments to estimate structural model parameters. See Altig, Christiano, Eichenbaum and Linde (2011), Christiano, Eichenbaum and Evans (2005) and Christiano, Eichenbaum and Trandbat (2013).

- If VARs are able to answer the questions above, then they may be useful to:
    - Provide guidance for alternative economic theories.
    - Or provide target strategies for estimating model parameters (e.g. matching an IRF).

- BUT, VARs cannot do any of that without help.
    - There is an IDENTIFICATION PROBLEM.
    - That can be, potentially, resolved with additional structure, i.e., theory (extra assumptions).

  It is with this additional structure that the VAR becomes a Structural VAR (SVAR).

The Mechanics of VARs

- Consider the following bivariate VAR representation

$$y_t = \mu + \Gamma y_{t-1} + \varepsilon_t \tag{16}$$

where $y_t = [y_{1,t}, y_{2,t}]'$, $\mu = [\mu_1, \mu_2]'$, Γ is a 2-by-2 matrix with generic element γ_{ij}, and $\varepsilon = [\varepsilon_1, \varepsilon_2]'$ with

$$\varepsilon_t \sim N(0, \Sigma),$$

that is, ε_1 and ε_2 are potentially correlated.

- For simplicity, we have assumed the existence of only one lag in (16).

- The system is also assumed to be covariance stationary.

- The *vector moving average* (VMA) representation (or Wold representation) of the VAR in (16) builds on an infinite moving average of innovations ε_t:[7]

$$y_t = \widetilde{\mu} + \Phi(L)\varepsilon_t$$

where

$$\Phi(L) = \sum_{k=0}^{\infty} \Phi_k L^k = \Phi_0 + \Phi_1 L + \Phi_2 L^2 + \Phi_3 L^3 + \dots$$

with $\Phi_k = \Gamma^k$ for $k > 0$ and $\Phi_0 = I$.[8]

- Invertibility conditions must hold.

[7] Where note that

$$\begin{aligned} y_t &= \mu + \Gamma y_{t-1} + \varepsilon_t \\ (1 - \Gamma L) y_t &= \mu + \varepsilon_t \\ y_t &= (1 - \Gamma L)^{-1} \mu + (1 - \Gamma L)^{-1} \varepsilon_t \\ y_t &= \widetilde{\mu} + \Phi(L) \varepsilon_t \end{aligned}$$

[8] Also, note that

$$\Phi(L) = \sum_{k=0}^{\infty} \Phi_k L^k = (1 - \Gamma L)^{-1} = I + \Gamma L + \Gamma^2 L^2 + \Gamma^3 L^3 + \dots$$

- Since the error terms ε_t are potentially contemporaneously correlated, in order to answer our original set of questions posed to the VAR we need to orthogonalize ε_t.

- We substitute $\varepsilon_t = \Omega u_t$, where $u_t \sim N(0, I)$.

- This will require some restriction to identify the elements in that we will discuss below. Theoretical restrictions will transform VARs into SVARs.

- The structural VMA representation is:

$$y_t = \widetilde{\mu} + \Theta(L)u_t$$

where $\Theta = \Phi(L)\Omega$ and,

$$\Theta(L) = \sum_{k=0}^{\infty} \Theta_k L^k = \Theta_0 + \Theta_1 L + \Theta_2 L^2 + \Theta_3 L^3 + \dots$$

with $\Theta_k = \Phi_k \Omega$ for all k. Note then that $\Theta_0 = \Phi_0 \Omega = I\Omega = \Omega$.

- Explicitly for the bivariate VAR case the VMA representation is

$$\begin{aligned}
\begin{bmatrix} y_{1,t} \\ y_{2,t} \end{bmatrix} &= \begin{bmatrix} \widetilde{\mu}_{1,t} \\ \widetilde{\mu}_{2,t} \end{bmatrix} + \Theta_0 \begin{bmatrix} u_{1,t} \\ u_{2,t} \end{bmatrix} + \Theta_1 L \begin{bmatrix} u_{1,t} \\ u_{2,t} \end{bmatrix} + \Theta_2 L^2 \begin{bmatrix} u_{1,t} \\ u_{2,t} \end{bmatrix} + \ldots \\
&= \begin{bmatrix} \widetilde{\mu}_{1,t} \\ \widetilde{\mu}_{2,t} \end{bmatrix} + \Theta_0 \begin{bmatrix} u_{1,t} \\ u_{2,t} \end{bmatrix} + \Theta_1 \begin{bmatrix} u_{1,t-1} \\ u_{2,t-1} \end{bmatrix} + \Theta_2 \begin{bmatrix} u_{1,t-2} \\ u_{2,t-2} \end{bmatrix} + \ldots
\end{aligned}$$

that is,

$$\begin{bmatrix} y_{1,t} \\ y_{2,t} \end{bmatrix} = \begin{bmatrix} \widetilde{\mu}_{1,t} \\ \widetilde{\mu}_{2,t} \end{bmatrix} + \begin{bmatrix} \theta_{11}^{(0)} & \theta_{12}^{(0)} \\ \theta_{21}^{(0)} & \theta_{22}^{(0)} \end{bmatrix} \begin{bmatrix} u_{1,t} \\ u_{2,t} \end{bmatrix} + \begin{bmatrix} \theta_{11}^{(1)} & \theta_{12}^{(1)} \\ \theta_{21}^{(1)} & \theta_{22}^{(1)} \end{bmatrix} \begin{bmatrix} u_{1,t-1} \\ u_{2,t-1} \end{bmatrix} + \ldots$$

- Then, at any time $t+s$ the structural VMA representation is

$$\begin{bmatrix} y_{1,t+s} \\ y_{2,t+s} \end{bmatrix} = \begin{bmatrix} \widetilde{\mu}_1 \\ \widetilde{\mu}_2 \end{bmatrix} + \Theta_0 \begin{bmatrix} u_{1,t+s} \\ u_{2,t+s} \end{bmatrix} + \Theta_1 L \begin{bmatrix} u_{1,t+s} \\ u_{2,t+s} \end{bmatrix} + \Theta_2 L^2 \begin{bmatrix} u_{1,t+s} \\ u_{2,t+s} \end{bmatrix} + \ldots$$
$$= \begin{bmatrix} \widetilde{\mu}_{1,t} \\ \widetilde{\mu}_{2,t} \end{bmatrix} + \Theta_0 \begin{bmatrix} u_{1,t+s} \\ u_{2,t+s} \end{bmatrix} + \ldots + \Theta_s \begin{bmatrix} u_{1,t} \\ u_{2,t} \end{bmatrix} + \ldots$$

that is,

$$\begin{bmatrix} y_{1,t+s} \\ y_{2,t+s} \end{bmatrix} = \begin{bmatrix} \widetilde{\mu}_{1,t} \\ \widetilde{\mu}_{2,t} \end{bmatrix} + \begin{bmatrix} \theta_{11}^{(0)} & \theta_{12}^{(0)} \\ \theta_{21}^{(0)} & \theta_{22}^{(0)} \end{bmatrix} \begin{bmatrix} u_{1,t+s} \\ u_{2,t+s} \end{bmatrix} + \ldots + \begin{bmatrix} \theta_{11}^{(s)} & \theta_{12}^{(s)} \\ \theta_{21}^{(s)} & \theta_{22}^{(s)} \end{bmatrix} \begin{bmatrix} u_{1,t} \\ u_{2,t} \end{bmatrix} + \ldots$$

- **Remark 1.** The effect of any innovation current u_t on any future y_t at period $t+s$ is given by Θ_s.

- **Remark 2.** The contemporaneous effect of any current innovation u_t on any current y_t is given by Θ_0.

- We define the *dynamic multipliers* as the response at $t+s$ of a variable y_i , $y_{i,t+s}$, to a shock u_j that happened at time t, $u_{j,t}$.

- The *dynamic multipliers* are

$$\frac{\partial y_{1t+s}}{\partial u_{1t}} = \theta_{11}^{(s)}, \quad \frac{\partial y_{1t+s}}{\partial u_{2t}} = \theta_{12}^{(s)}$$

$$\frac{\partial y_{2t+s}}{\partial u_{1t}} = \theta_{21}^{(s)}, \quad \frac{\partial y_{2t+s}}{\partial u_{2t}} = \theta_{22}^{(s)},$$

Impulse Response Functions

- IRFs describe the response at each period $t + s$, for all $s > 0$, of variable y_{it+s} to a shock u_{jt} that happens at time t.

- That is, IRFs are defined by the sequence of dynamic multipliers $\left\{\theta_{ij}^{(s)}\right\}$.

What is the impact in the long-run?

- Since y_t is assumed covariance stationary, then the limiting dynamic multiplier is

$$\lim_{s\to\infty} \frac{\partial y_{it+s}}{\partial u_{jt}} = \lim_{s\to\infty} \theta_{ij}^{(s)} = 0$$

- That is, shocks to u_j have no long-run impact on the level of y_i. In other terms, the response of y_{it+s} in the long-run to shocks u_{jt} at time t is zero. The reason is that y_t (we have assumed) is covariance stationary.

What is the cumulative impact in the long-run?

- The cumulative impact of u_j on y_i in the long-run is given by

$$\Theta(1) = \begin{bmatrix} \theta_{11}(1) & \theta_{12}(1) \\ \theta_{21}(1) & \theta_{22}(1) \end{bmatrix} = \begin{bmatrix} \sum_{s=0}^{\infty} \theta_{11}^{(s)} & \sum_{s=0}^{\infty} \theta_{12}^{(s)} \\ \sum_{s=0}^{\infty} \theta_{21}^{(s)} & \sum_{s=0}^{\infty} \theta_{22}^{(s)} \end{bmatrix}$$

where note that, $\Theta(1) = \sum_{k=0}^{\infty} \Theta_k (L = 1)^k = \Theta_0 + \Theta_1 + \Theta_2 + \Theta_3 + \dots$

- That is, the cumulative impact of u_j on y_i in the long-run is the infinite sum of dynamic multipliers.

Forecast Error

- Forecast error is the difference between the actual value of y_{t+s},

$$\begin{aligned} y_{t+s} &= \widetilde{\mu} + \Phi(L)\varepsilon_{t+s} \\ &= \widetilde{\mu} + \Phi_0 L^0 \varepsilon_{t+s} + \Phi_1 L^1 \varepsilon_{t+s} + \ldots + \Phi_{s-1} L^{s-1} \varepsilon_{t+s} + \Phi_s L^s \varepsilon_{t+s} + \Phi_{s+1} L^{s+1} \varepsilon_{t+s} + \ldots \\ &= \widetilde{\mu} + \Phi_0 \varepsilon_{t+s} + \Phi_1 \varepsilon_{t+s-1} + \ldots + \Phi_{s-1} \varepsilon_{t+1} + \Phi_s \varepsilon_t + \Phi_{s+1} \varepsilon_{t-1} + \ldots \end{aligned}$$

and the predicted value for $\widehat{y}_{t+s}$ given information available at time t,

$$\widehat{y}_{t+s|t} = \widetilde{\mu} + \Phi_s \varepsilon_t + \Phi_{s+1} \varepsilon_{t-1} + \ldots$$

That is, **forecast error** is

$$y_{t+s} - \widehat{y}_{t+s|t} = \Phi_0 \varepsilon_{t+s} + \Phi_1 \varepsilon_{t+s-1} + \ldots + \Phi_{s-1} \varepsilon_{t+1}$$

- Using $\varepsilon_t = \Omega u_t$, forecast error is

$$y_{t+s} - \widehat{y}_{t+s|t} = \Theta_0 u_{t+s} + \Theta_1 u_{t+s-1} + \ldots + \Theta_{s-1} u_{t+1}$$

that is,

$$\begin{bmatrix} y_{1,t+s} - \widehat{y}_{1,t+s|t} \\ y_{2,t+s} - \widehat{y}_{2,t+s|t} \end{bmatrix} = \begin{bmatrix} \theta_{11}^{(0)} & \theta_{12}^{(0)} \\ \theta_{21}^{(0)} & \theta_{22}^{(0)} \end{bmatrix} \begin{bmatrix} u_{1,t+s} \\ u_{2,t+s} \end{bmatrix} + \ldots + \begin{bmatrix} \theta_{11}^{(s-1)} & \theta_{12}^{(s-1)} \\ \theta_{21}^{(s-1)} & \theta_{22}^{(s-1)} \end{bmatrix} \begin{bmatrix} u_{1,t+1} \\ u_{2,t+1} \end{bmatrix}$$

- Let's focus first on the first row...

- The forecast error associated with y_1 is

$$y_{1t+s} - \widehat{y}_{1t+s|t} = \theta_{11}^{(0)} u_{1t+s} + \ldots + \theta_{11}^{(s-1)} u_{1t+1} + \theta_{12}^{(0)} u_{2t+s} + \ldots + \theta_{12}^{(s-1)} u_{2t+1}$$

- Because u_{1t} and u_{2t} are not related to each other, and not autocorrelated, then $Var(u_i) = Var(u_{it}) = Var(u_{it+s})$ for any s, hence, the variance of the forecast error is

$$\begin{aligned} Var(y_{1t+s} - \widehat{y}_{1t+s|t}) =& Var(u_1)\left(\left(\theta_{11}^{(0)}\right)^2 + \ldots + \left(\theta_{11}^{(s-1)}\right)^2\right) + \\ & Var(u_2)\left(\left(\theta_{12}^{(0)}\right)^2 + \ldots + \left(\theta_{12}^{(s-1)}\right)^2\right) \end{aligned}$$

Forecast Error Variance Decomposition

- The forecast error of y_1 generated by u_1 is

$$\sigma^2_{y_1}(u_1) = \frac{Var(u_1)\left(\left(\theta_{11}^{(0)}\right)^2 + \ldots + \left(\theta_{11}^{(s-1)}\right)^2\right)}{Var(y_{1t+s} - \widehat{y}_{1t+s|t})}$$

- The forecast error of y_1 generated by u_2 is

$$\sigma^2_{y_1}(u_2) = \frac{Var(u_2)\left(\left(\theta_{12}^{(0)}\right)^2 + \ldots + \left(\theta_{12}^{(s-1)}\right)^2\right)}{Var(y_{1t+s} - \widehat{y}_{1t+s|t})}$$

- The forecast errors associated with y_2 are analogous.

Implementation of Short-Run Restrictions

- Short-run restrictions are associated with contemporaneous effects and, therefore, are implemented on Θ_0.

- For example, the restriction imposed by

$$\Theta_0 = \begin{bmatrix} \theta_{11}^{(0)} & 0 \\ \theta_{21}^{(0)} & \theta_{22}^{(0)} \end{bmatrix}$$

implies that u_{2t} has no contemporaneous effect on y_{1t} , that is, $\theta_{12}^{(0)} = 0$.

Implementation of Long-Run Restrictions

- Long-run restrictions are associated with the cumulative long-run effects and, therefore, are implemented on $\Theta(1)$.

- For example, the restriction imposed by

$$\Theta(1) = \begin{bmatrix} \theta_{11}(1) & 0 \\ \theta_{21}(1) & \theta_{22}(1) \end{bmatrix} = \begin{bmatrix} \sum_{s=0}^{\infty} \theta_{11}^{(s)} & 0 \\ \sum_{s=0}^{\infty} \theta_{21}^{(s)} & \sum_{s=0}^{\infty} \theta_{22}^{(s)} \end{bmatrix}$$

  implies that u_{2t} has no cumulative long-run effect on y_{1t}, that is, $\theta_{12}(1) = \sum_{s=0}^{\infty} \theta_{12}^{(s)} = 0$.

Short-Run Identification of SVARs: Example 1: Identification of Labor-share shocks and Productivity Shocks

- Working example: Ríos-Rull and Santaeulàlia-Llopis (2010).

- We are interested in 2 stationary series: linearly detrended productivity residual, z_t^1; and labor share (deviations from mean), z_t^2.

- The productivity residual is recovered from data on output, hours, capital and labor share.

- We may be interested not only in the cyclical properties of z_t^1 and z_t^2, but also in the systematic joint behavior between z_t^1 and z_t^2.

- To get a better idea of this joint dynamics we can use a vector AR system, and compute IRFs and FEVDs.

- We now pose a statistical model to find an underlying stochastic process that generates the joint distribution of z_t^1 and z_t^2.

- That is,

$$z_t = \Gamma z_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim N(0, \Sigma),$$

  where $z_t = (z_t^1, z_t^2)'$ and Γ is a 2-by-2 square matrix with generic element γ_{ij}. The innovations $\varepsilon_t = (\varepsilon_t^1, \varepsilon_t^2)'$ are serially uncorrelated and follow a bivariate Gaussian distribution with unconditional mean zero and a symmetric positive definite variance-covariance matrix Σ.

- Thus, this specification has seven parameters: the four coefficient regressors in Γ, and the variances and covariance in Σ.

Maximum Likelihood may yield the following estimates,

$$\widehat{\Gamma} = \begin{pmatrix} .946 & .001 \\ (.023) & (.042) \\ .050 & .930 \\ (.010) & (.019) \end{pmatrix}, \qquad \widehat{\Sigma} = \begin{pmatrix} .00668^2 & -.1045E-04 \\ -.1045E-04 & .00304^2 \end{pmatrix}.$$

This generates a negative contemporaneous correlation between innovations ε_t of -.51.

- Further, since our innovations ε_t are contemporaneously correlated, we would like to transform ε_t to a set of uncorrelated components u_t according to $\varepsilon_t = \Omega\ u_t$, where Ω is an invertible square matrix with generic element ω_{ij}, such that

$$\widehat{\Sigma} = \frac{1}{n} \sum_t \varepsilon_t \varepsilon_t^{'} = \Omega \left(\frac{1}{n} \sum_t u_t u_t^{'} \right) \Omega^{'} = \Omega\Omega'$$

  and we have normalized u_t to have unit variance.

- And here it is the problem...

Identification Problem

- While $\widehat{\Sigma}$ has three parameters, the matrix Ω has four: there are many such matrices.
- To see this note that $\widehat{\sigma} = \begin{bmatrix} \widehat{\sigma}_{11} & \widehat{\sigma}_{12} \\ \widehat{\sigma}_{21} & \widehat{\sigma}_{22} \end{bmatrix}$ where by symmetry we have that

  $\widehat{\sigma}_{12} = \widehat{\sigma}_{21}$ and we have to solve for the elements in $\Omega = \begin{bmatrix} \omega_{11} & \omega_{12} \\ \omega_{21} & \omega_{22} \end{bmatrix}$ such that $\widehat{\Sigma} = \Omega\Omega'$.
- This means that we have 4 unknowns in Ω but only three restrictions coming from $\widehat{\Sigma} = \Omega\Omega'$ and hence we cannot solve for the elements in Ω in a unique manner unless we impose (at least) on further restriction.
- To explicitly see this identification problem we can expand $\widehat{\Sigma} = \Omega\Omega'$ to find a four-equations system:

  $\widehat{\sigma}_{11} = \omega_{11}^2 + \omega_{12}^2,\ \widehat{\sigma}_{12} = \omega_{11}\omega_{21} + \omega_{12}\omega_{22},\ \widehat{\sigma}_{21} = \omega_{11}\omega_{21} + \omega_{12}\omega_{22},$ and $\widehat{\sigma}_{22} = \omega_{21}^2 + \omega_{22}^2$

  Since $\widehat{\sigma}_{12} = \widehat{\sigma}_{21}$, the system with 4 equations above is actually a system with three equations:

  $$\widehat{\sigma}_{11} = \omega_{11}^2 + \omega_{12}^2,\ \widehat{\sigma}_{12} = \omega_{11}\omega_{21} + \omega_{12}\omega_{22},\ \text{and } \widehat{\sigma}_{22} = \omega_{21}^2 + \omega_{22}^2$$

  and 4 unkowns, $\omega_{11}, \omega_{12}, \omega_{21}$, and ω_{22}. This implies that there are infinite solutions to the problem.

Identification Strategy

- A solution to the identification problem: Implement a short-run restriction:

  Identification assumption. [Ríos-Rull and Santaeulàlia-Llopis (2010)] *Only productivity innovations, $u_{z,t}$, have a short-run effect on productivity.*

- What the identification assumption implies is that innovations to factor shares are purely redistributive in nature: without contemporaenous effects on productivity.

- To see this recall that the contamparenous, i.e., short-run, effects are captured by Θ_0, and note that $\Theta_0 = \Omega$

- Explicitly, the short-run restriction implied by the identification assumption is the setting of Ω to be a lower triangular matrix. That is, since $\Theta_0 = \Omega$, the short-run restriction that we are imposing is

$$\theta_{12}^{(0)} = \omega_{12} = 0$$

  One way to implement this restriction is through a Cholesky factorization of $\widehat{\Sigma}$.

- Our Cholesky factorization of $\widehat{\Sigma}$ results in

$$\begin{pmatrix} \varepsilon_t^1 \\ \varepsilon_t^2 \end{pmatrix} = \begin{pmatrix} \omega_{11} & \omega_{12} \\ \omega_{12} & \omega_{22} \end{pmatrix} \begin{pmatrix} u_t^1 \\ u_t^2 \end{pmatrix} = \begin{pmatrix} .00668 & .0 \\ -.00156 & .00260 \end{pmatrix} \begin{pmatrix} u_t^1 \\ u_t^2 \end{pmatrix},$$

where $\omega_{11} = \sigma_{\varepsilon^1}$, $\omega_{21} = E[\varepsilon_t^2 | \varepsilon_t^1]$, and ω_{22} is the standard error of the regression of ε_t^2 on ε_t^1.

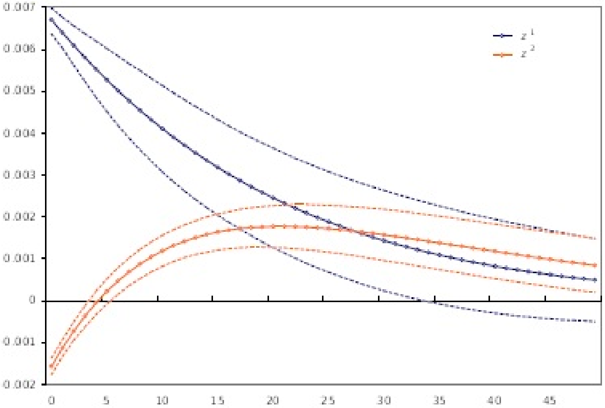


Figure: IRFs to Orthogonalized Productivity Innovations u^1

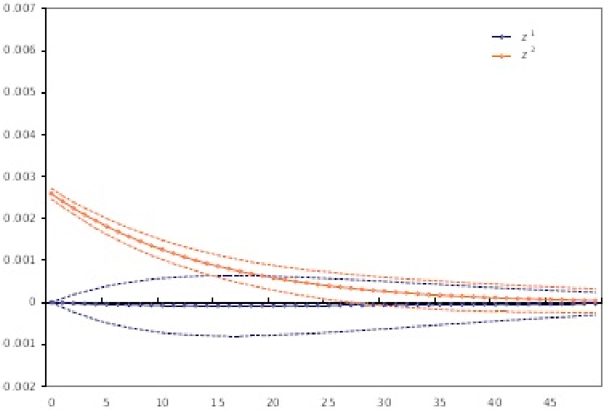


Figure: IRFs to Orthogonalized Redistributive Innovations u^2

Long-Run Identification of SVARs: Example 2: Identification of Productivity and Investment Shocks

- Recently, there has been a heated empirical debate on the contribution of productivity shocks on hours (and output) that results from long-run restrictions applied to SVARs.
- Galí (1999) and Fisher (2006) are exponents of that debate and we will use them as working examples.

- The original argument by Galí (1999) is that, in many business cycle modelizations, only productivity shocks have a long-run impact on labor productivity, lp_t.

- This is the case of the standard RBC model where with only productivity shocks (as in Prescott (1986) and Cooley and Prescott (1995)) labor productivity has a trend solely determined by Z_t. Therefore, if Z_t has a stochastic trend, then so does labor productivity (see, e.g., King et al. (1988), Hansen (1997), and King and Rebelo (1999));

- Hence, in principle, one could use that theoretical bullet to run a bivariate VAR on on labor productivity and (say) hours per capita; impose the long-run restriction that the shocks to hours have no long-run effect on labor productivity, hence, identify productivity shocks as the only ones that have long-run impact on labor productivity; and, then evaluate the contribution of productivity shocks, identified in such manner, to the fluctuation of hours.

- We will conduct this exercise next, but first we discuss the source of identification: a standard business cycle theory with productivity and investment shocks, Fisher (2006).

A Standard Business Cycle Model with Productivity and Investment Shocks, Fisher (2006)

Consider a social planner that maximizes the discounted value of future utility

$$E_0 \sum_{t}^{\infty} \beta^t N_t u(c_t, h_t) \tag{17}$$

by choosing stochastic streams of consumption per capita and hours per capita, $\{c_t, h_t\}_{t=0}^{\infty}$, with discount factor $\beta \in (0, 1)$, and population level N_t growing at a constant rate λ_N, that is, $N_t = N_0 \, e^{\lambda_N t}$. E_0 is the conditional expectation operator.

The problem is subject to a resource constraint is,

$$C + \frac{1}{V_t} X_t = Y_t, \tag{18}$$

where $\frac{1}{V_t} X_t = I_t$, X_t is investment in efficiency units and ISTC, V_t, is the inverse of the relative price of efficiency investment units in terms of consumption units, P_t^I, i.e., $V_t = \frac{1}{P_t^I}$. Note that, given the resource constraint (18), it is the case that in the long run $\lambda_C = \lambda_I = \lambda_Y$.

Further, the law of motion of capital is,

$$X_t = V_t \; I_t = K_{t+1} - (1-\delta)K_t \tag{19}$$

This implies that for capital to grow at a constant rate in the long run the limiting investment to capital ratio in efficiency units, i.e., $\lim_{s\to\infty} \frac{V_{t+s} I_{t+s}}{K_{t+s}}$, must be constant; that is, $\lambda_V + \lambda_I = \lambda_K$.

Aggregate output, Y_t, is produced with a general constant returns to scale (CRS) production function,

$$Y_t = A_t\ F(H_t, K_t), \tag{20}$$

where A_t is neutral technical change and $H_t = N_t h_t$ is aggregate hours. Initial conditions $A_0,\ K_0$ and N_0 are given.

Neutral technical change is exogenous and follows a stochastic trend,

$$A_t = A_{t-1}\ e^{\lambda_A + \varepsilon_{A,t}}, \tag{21}$$

that is, the log of A_t follows a random walk with drift λ_A, where λ_A is the long-run growth rate of A_t and a source of fluctuations in this model are productivity shocks, $\varepsilon_{A,t}$.

Finally, we assume that ISTC is exogenously given and follows

$$V_t = V_{t-1}\ e^{\lambda_V + \varepsilon_{V,t}}, \tag{22}$$

that is, the log of V_t follows a random walk with drift λ_V, where λ_V is the long-run growth rate of ISTC and investment shocks $\varepsilon_{V,t}$ are a source of fluctuations.

We next derive the long-run implications of productivity shocks and investment shocks on labor productivity in this model. To see this, we totally differentiate the left and right hand sides of the aggregate production function (20) with respect to time and divide by Y_t to find that the growth rate of output in the long-run is $\lambda_Y = \lambda_A + ls\lambda_H + (1 - ls)\lambda_K$.

Then, after rearranging, and using the fact that $\lambda_Y = \lambda_K - \lambda_V$, the long-run growth rate of labor productivity becomes

$$\lambda_{lp} = \frac{1}{ls}\lambda_A + \frac{1 - ls}{ls}\lambda_V. \tag{23}$$

That is, long-run labor productivity grows due to either neutral technical change A_t (up to a constant $\frac{1}{ls}$) or investment-specific technical change V_t (up to a constant $\frac{1-ls}{ls}$). Therefore, if ls_t is stationary (i.e., $\lim_{s\to\infty} ls_{t+s} = ls$) then only productivity shocks $\varepsilon_{A,t}$ and investment shocks $\varepsilon_{V,t}$ have a permanent effect on the level of labor productivity. Specifically, the long-run relationship between lp_t and V_t implied by (23) is

$$\lim_{s\to\infty} \frac{\ln lp_{t+s}}{\epsilon_{V,t}} = \frac{1-ls}{ls} \lim_{s\to\infty} \frac{\ln V_{t+s}}{\epsilon_{V,t}} = \frac{1-ls}{ls}, \tag{24}$$

and between lp_t and A_t,

$$\lim_{s\to\infty} \frac{\ln lp_{t+s}}{\epsilon_{A,t}} = \frac{1}{ls} \lim_{s\to\infty} \frac{\ln A_{t+s}}{\epsilon_{A,t}} = \frac{1}{ls}. \tag{25}$$

Finally, the fact that in this model the real price of investment in terms of consumption is $P_t^I = \frac{1}{V_t}$ implies that

$$\lim_{s\to\infty} \frac{\ln P_{t+s}}{\epsilon_{V,t}} = -\lim_{s\to\infty} \frac{\ln V_{t+s}}{\epsilon_{V,t}} = -1. \tag{26}$$

That is, only investment shocks have a permanent effect on the relative price of investment. We can summarize these theoretical implications in the following result.

Theoretical Result. [Standard RBC model with ISTC] *In a model described by the maximization of (17) subject to resource constraint (18), capital law of motion (19), aggregate production function (20), exogenous neutral technical change (21), and investment-specific technical change (22), productivity shocks and investment shocks are the only ones that have a permanent effect on the level of labor productivity; and, investment shocks are the only ones that have a permanent effect on the real price of investment.*

The Theoretical Result informs the identification strategy of productivity shocks and investment shocks proposed by Fisher (2006) and that we implemented next.

Under the assumption of a Cobb-Douglas production function with constant coefficients, the following trivariate VAR system helps assess the contribution of productivity shocks, A_t, and investment shocks, V_t, to aggregate fluctuations. Pose

$$\begin{pmatrix} \lambda_{v,t} \\ \lambda_{lp,t} \\ h_t \end{pmatrix} = \begin{pmatrix} \phi_{\lambda_v,\lambda_v}(L) & \phi_{\lambda_v,\lambda_{lp}}(L) & \phi_{\lambda_v,h}(L) \\ \phi_{\lambda_{lp},v}(L) & \phi_{\lambda_{lp},\lambda_{lp}}(L) & \phi_{\lambda_{lp},h}(L) \\ \phi_{h,\lambda_v}(L) & \phi_{h,\lambda_{lp}}(L) & \phi_{h,h}(L) \end{pmatrix} \begin{pmatrix} \varepsilon_{\lambda_v,t} \\ \varepsilon_{\lambda_{lp},t} \\ \varepsilon_{h,t} \end{pmatrix}, \quad (27)$$

where λ_{v_t} is the growth rate of the inverse of the relative price of investment, λ_{lp_t} is the growth rate of labor productivity, h_t are hours per capita (in log-levels), and $\varepsilon_{\lambda_v,t}$, $\varepsilon_{\lambda_{lp},t}$, and $\varepsilon_{h,t}$ are contemporaneously related with a 3×3 covariance-variance matrix Σ.

The three endogenous series are stationary, and we can estimate $\widehat{\phi}_{ij}(L)$ and $\widehat{\Sigma}$ with standard techniques.

Identification problem

Since $\varepsilon_{\lambda_v,t}$, $\varepsilon_{\lambda_{lp},t}$, and $\varepsilon_{h,t}$ are contemporaneously correlated, assessing the contribution of each shock to aggregate fluctuations requires transforming ε_t into orthogonal innovations $u_t = \Omega^{-1}\varepsilon_t$, where Ω is an invertible square matrix with $\widehat{\Sigma} = \Omega\Omega'$. Further, the variance of u_t is normalized to 1.

An identification problem arises because while Ω has nine parameters, the restriction $\widehat{\Sigma} = \Omega\Omega'$ poses only six equations. We need (at least) three additional.

Identification strategy

We resolve this identification problem with three long-run restrictions informed by our Theoretical Result. To this end, the VMA system (27) can be written in terms of the orthogonal innovations, u_t,

$$\begin{pmatrix} \lambda_{v,t} \\ \lambda_{lp,t} \\ h_t \end{pmatrix} = \begin{pmatrix} \theta_{\lambda_v,\lambda_v}(L) & \theta_{\lambda_v,\lambda_{lp}}(L) & \theta_{\lambda_v,h}(L) \\ \theta_{\lambda_{lp},v}(L) & \theta_{\lambda_{lp},\lambda_{lp}}(L) & \theta_{\lambda_{lp},h}(L) \\ \theta_{h,\lambda_v}(L) & \theta_{h,\lambda_{lp}}(L) & \theta_{h,h}(L) \end{pmatrix} \begin{pmatrix} u_{\lambda_v,t} \\ u_{\lambda_{lp},t} \\ u_{h,t} \end{pmatrix}, \quad (28)$$

where $\Theta(L) = \Phi(L)\Omega$; and $u_{\lambda_v,t}$, $u_{\lambda_{lp},t}$, and $u_{h,t}$ are orthogonal with variance 1.

Identification assumption [Fisher (2006)]

1. *Only investment innovations,* $u_{\lambda_v,t}$*, have a long-run effect on ISTC,* V_t*.*

2. *Only investment innovations,* $u_{\lambda_v,t}$*, and productivity innovations,* $u_{\lambda_{lp},t}$*, have a long-run effect on labor productivity,* lp_t*.*

This identification assumption implies three long-run restrictions. Specifically, Assumption C1 implies the following two restrictions:

$$\lim_{s\to\infty} \frac{\partial V_{t+s}}{\partial u_{lp,t}} = \theta_{\lambda_V,\lambda_{lp}}(1) = \sum_{s=0}^{\infty} \theta^{(s)}_{\lambda_V,\lambda_{lp}} = 0,$$

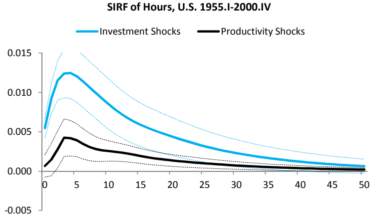
$$\lim_{s\to\infty} \frac{\partial V_{t+s}}{\partial u_{h,t}} = \theta_{\lambda_V,h}(1) = \sum_{s=0}^{\infty} \theta^{(s)}_{\lambda_V,h} = 0,$$

and Assumption C2 implies one restriction:

$$\lim_{s\to\infty} \frac{\partial lp_{t+s}}{\partial u_{h,t}} = \theta_{\lambda_{lp},h}(1) = \sum_{s=0}^{\infty} \theta^{(s)}_{\lambda_{lp},h} = 0.$$

Thus, the trivariate VAR system (28) and Assumption C structurally identify all elements in Ω, as in Fisher (2006).

Results



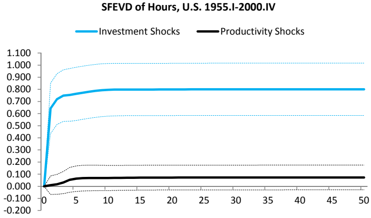


Figure: SIRFs and FEVDs with Long-Run Restrictions, Fisher (2006) Model, U.S. 1955.I-2000.IV; 68% bootstrap error bands; Author's calculations.

Summary of results

- Investment shocks contribute largely to the variance of hours, 79%.
- The contribution of productivity shocks is small, around 7%.

Note that Fisher (2006) splits the sample into two periods 1955-1979, and 1982-2000; see a discussion on alternative trend breaks in Fernald (2007), Canova et al. (2010), and Santaeulàlia-Llopis (2012). Here, for simplicity, I have ignored the possibility of trend breaks.

- The 'advantage' of this approach is that we do not need extra assumptions required to construct the productivity residual, in particular, corrections for labor hoarding, capital utilizaiton and/or time-varying markups.

- The 'disadvantage' of these approach is that some models may not satisfy the long-run identifying assumption. In particular, endogenous growth models have all shocks affecting productivity in the long-run; and permanent shocks to tax rates on capital income also have long-run effect on productivity.

Long-Run Identification of SVARs: Example 3: Identification of Spillover Effects in China

- An economy with 3 sectors for China: Foreign-invested enterprises (FIE), private sector (PRI), and state-owned enterprises (SOE).
- Our theory implies that there are long-run effects from FIE innovations to PRI and SOE productivity, and from PRI innovations to SOE productivity. That is, this theory has the following two long-run implications:

  1. Only innovations to productivity in FIE affect the long-run value of productivity in FIE.
  2. Only innovations to productivity in FIE, and innovations to productivity in PRI, affect the long-run value of productivity in PRI.

  That is, theory implies three long-run restrictions. Next we apply these restrictions to a trivariate VAR system using data on labor productivity growth in the three sectors: FIE, PRI, and SOE in China 1996-2010.[9]

[9] Source of data: CEIC China Premium Database. These aggregate variables by type of industrial enterprises are obtained from the section of Industry Overview and Financial Data. The output series is the value added; the labor input is the number of employees.

	FIE	PRI	SOE
FIE Productivity Innovations	.78	.55	.05
	[.52,1.05]	[.36,.74]	[-.12,.22]
PRI Productivity Innovations	.05	.36	.55
	[-.14,.24]	[.21,.52]	[.32,.77]
SOE Productivity Innovations	.17	.09	.40
	[-.03,.36]	[-.07,.24]	[.21,.59]

Table: SFEVDs: Contribution of productivity innovations of each sector (FIE, PRI, and SOE) to labor productivity growth in each sector; China 1996-2010

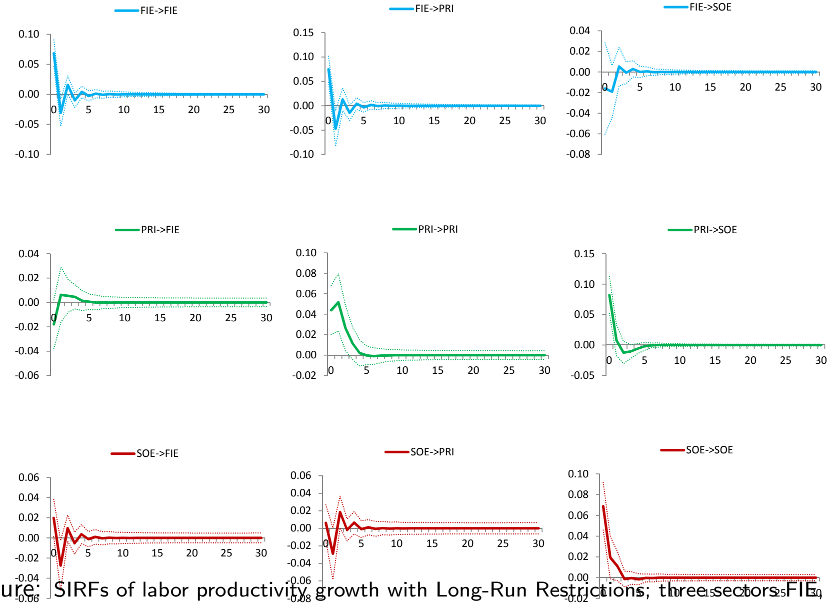


Figure: SIRFs of labor productivity growth with Long-Run Restrictions; three sectors FIE, PRI, and SOE in China 1996-2010; 68% bootstrap error bands; Author's calculations.

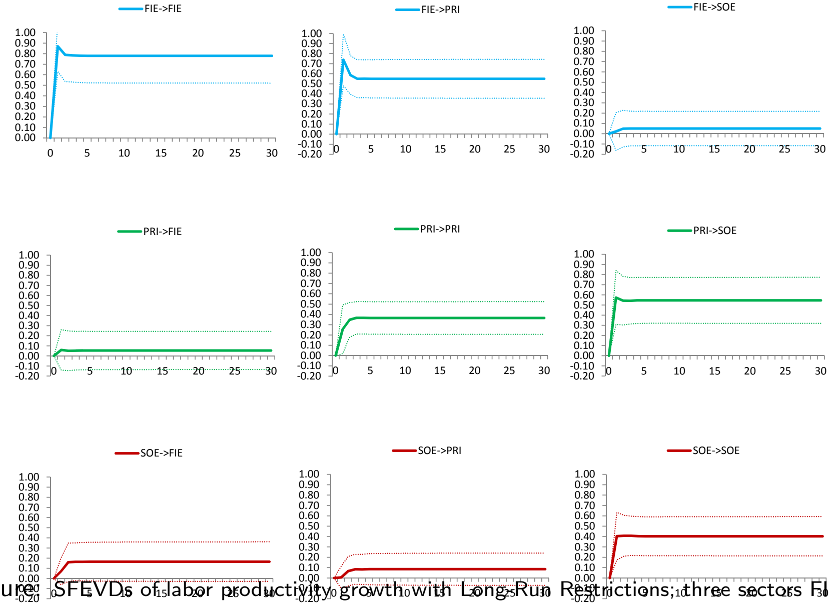


Figure: SFEVDs of labor productivity growth with Long-Run Restrictions; three sectors FIE, PRI, and SOE in China 1996-2010; 68% bootstrap error bands; Author's calculations.

Summary of results

- Technological innovations in FIE account for almost all its own labor produtivity growth, 78%.
- The evidence of spillover effects is clear: There are large and significant effects of FIE technological innovations on PRI labor productivity growth, about 55%.
- Further, there are also large and significant effects of PRI technological innovations on SOE labor productivity growth, about 55%.

Bai, J. and Perron, P. (1998). Testing for and estimation of multiple structural changes. *Econometrica*, 66(1):47–78.

Canova, F., López-Salido, D., and Michelacci, C. (2010). On the robustness effects of technology shocks on hours worked and output. *Journal of Applied Econometircs*. Forthcoming.

Chari, V. V., Kehoe, P. J., and McGrattan, E. (2008). Are structural vars with long-run restrictions useful in developing business cycle theory? *Journal of Monetary Economics*, 55(8):1337–1352.

Christiano, L. J. and Eichenbaum, M. (1988). Unit roots in real gnp: Do we know and do we care? Forthcoming in *Carnegie-Rohester Conference Series on Public Policy* 32. Amsterdam, North Holland.

Christiano, L. J., Eichenbaum, M., and Vigfusson, R. (2006). Assessing structural vars. In Acemoglu, D., Rogoff, K., and Woodford, M., editors, *NBER Macroeconomics Annual 2006*. MIT Press.

Cooley, T. F. and Prescott, E. C. (1995). Economic growth and business cycles. In Cooley, T. F., editor, *Frontiers of Business Cycle Research*, pages 1–38. Princeton University Press, Princeton, NJ.

Cummins, J. and Violante, G. L. (2002). Investment-specific technical change in the us (1947-2000): Measurement and macroeconomic consequences. *Review of Economic Dynamics*, 5:243–284.

Fernald, J. G. (2007). Trend breaks, long-run restrictions, and contractionary technology improvements. *Journal of Monetary Economics*, 54(8):2467–2485.

Fisher, J. D. (2006). The dynamic effects of neutral and investment-specific technology shocks. *Journal of Political Economy*, 114(3):413–451.

Galí, J. (1999). Technology, employment, and the business cycle: Do technology shocks explain aggregate fluctuations? *American Economic Review*, 89(1):249–271.

Gordon, R. J. (1990). *The Measurement of Durable Goods Prices*. National Bureau of Economic Research Monograph Series, Chicago: University of Chicago Press.

Hansen, G. D. (1997). Technical progress and aggregate fluctuations. *Journal of Economic Dynamics and Control*, 21(6):1005–1023.

King, R. G., Plosser, C. I., and Rebelo, S. (1988). Production, growth and business cycles: Ii new directions. *Journal of Monetary Economics*, 21:309–41.

King, R. G. and Rebelo, S. T. (1999). Resuscitating real business cycles. In Taylor, J. B. and Woodford, M., editors, *Handbook of Macroeconomics*, chapter 14, pages 927–1007. Elsevier, Amsterdam.

Perron, P. (1990). Testing for a unit root in a time series with a changing mean. *Journal of Business and Economic Statistics*, 8(2):153–162.

Prescott, E. C. (1986). Theory ahead of business-cycle measurement. 25(1):11–44. Also in *Federal Reserve Bank of Minneapolis Quarterly Review*, 10(4), 9–22.

Ravn, M. O. and Uhlig, H. (2002). On adjusting the hodrick-prescott filter for the frequency of observations. *The Review of Economic and Statistics*, 84(2):371–376.

Ríos-Rull, J.-V. and Santaeulàlia-Llopis, R. (2010). Redistributive shocks and productivity shocks. *Journal of Monetary Economics*, 57(8):931–948.

Ríos-Rull, J.-V., Schorfheide, F., Fuentes-Albero, C., Kryshko, M., and Santaeulàlia-Llopis, R. (2012). Methods versus substance: Measuring the effects of technology shocks. *Journal of Monetary Economics*, 57(8):931–948.

Santaeulàlia-Llopis, R. (2012). How much are svars with long-run restrictions missing without cyclically moving factor shares? Mimeo, Washington University in St. Louis.