# *Income Inequality*

ECON 815
University of South Carolina

Fall 2017

# Defining terms

- What is income inequality?
  - What is *income*?
  - What is *inequality*?

# What is income?

- Is your paycheck income?
- Suppose you own Facebook stock - what if the stock increases in value - is that income?
- This is actually a deep question.
- Even accountants disagree:
    - Cash vs accrual accounting?

# Economic income

- Economic income is called "Haig-Simons income"
- Definition:
  - Income is anything that *changes* your ability to *consume*
- Note *changes*, implying that income is a *flow* (wealth is a stock)
- Note the emphasis on consumption
  - Utility from consumption
  - Not Scrooge McDuck utility

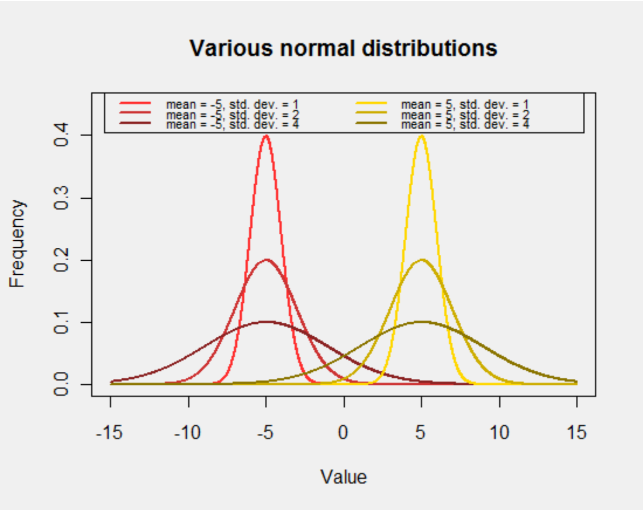
# Measuring income in practice

- Not Haig-Simons income
- More like cash accounting (though sometimes include value of non-cash benefits)
- Unit of observation
  - Individual?
  - Tax unit?
  - Household?
- Data
  - Survey (CPS, PSID, SCF, others)
  - Administrative (SSA, IRS)
  - **Important**
    - Panel data vs. cross-sectional data
    - What is measured in income?
    - Can you observe income of the very top earners?

# What is inequality?

- An unequal *distribution*
- So what's a distribution?
- And how do we measure the degree to which it is "unequal"?

# Distributions




One measure of the "unequal-ness" of the distribution is the variance/standard deviation

# Measures of inequality

- Cross-sectional variance
- Gini coefficient
- Percentile ratios (e.g., the income of the 90th percentile divided by the income of the 10th percentile)
- Fraction of total income by percentile (e.g., how much of total income goes to the top 1%)

# U.S. income inequality facts

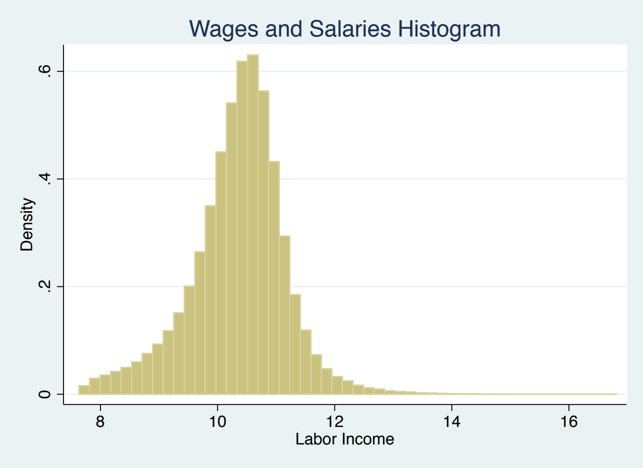
Pretty much all data/measures say the same things (though the magnitudes differ):

- Income inequality is increasing
- It has been increasing since at least since the 1970s.
- The increase in inequality is largely driven by increases in income at the very top of the distribution

This is true in:

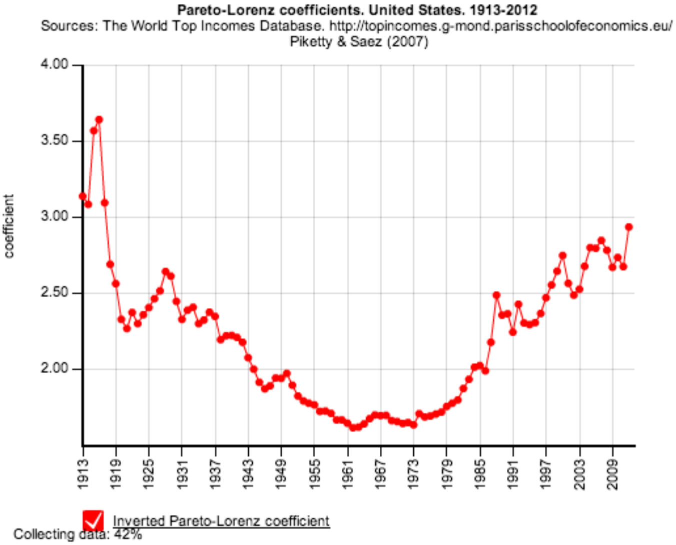
- Panel and cross-sectional data
- Survey and administrative data
- For most measures of income (e.g. earned and unearned income)

# The distribution of Income in the U.S.

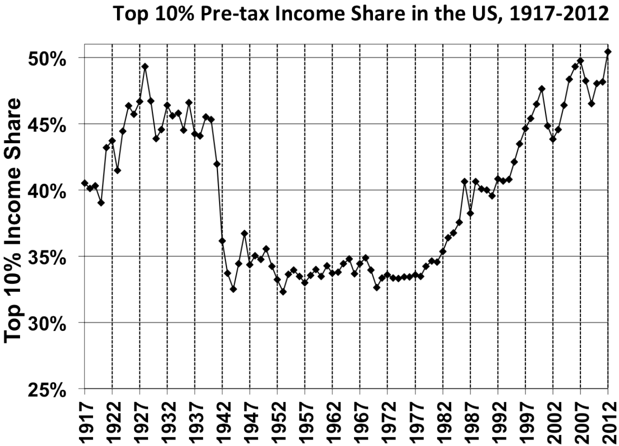



Wages and salaries are distributed approximately log-normally

# The income distribution over time (1)

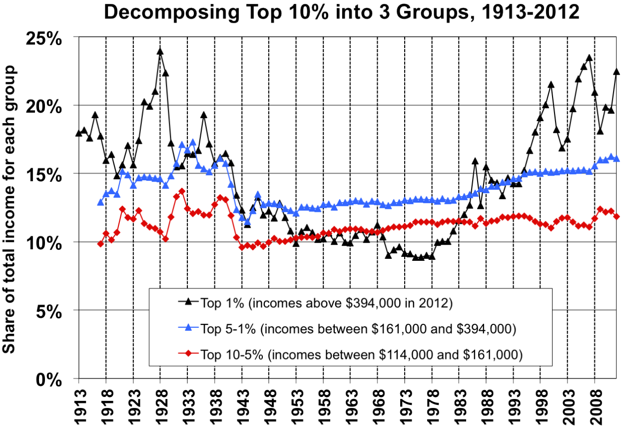


# The income distribution over time (2)



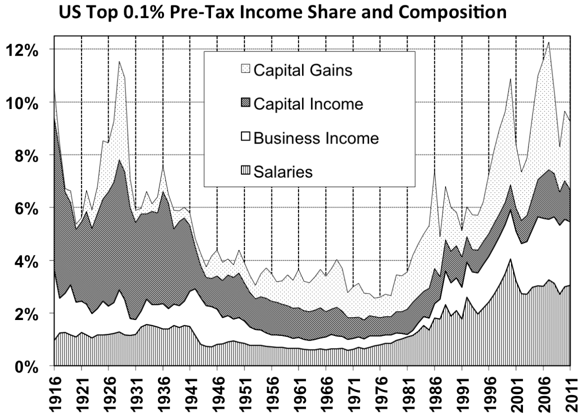

Source: Piketty and Saez, 2003 updated to 2012. Series based on pre-tax cash market income including realized capital gains and excluding government transfers. 2012 data based on preliminary statistics

# The income distribution over time (3)



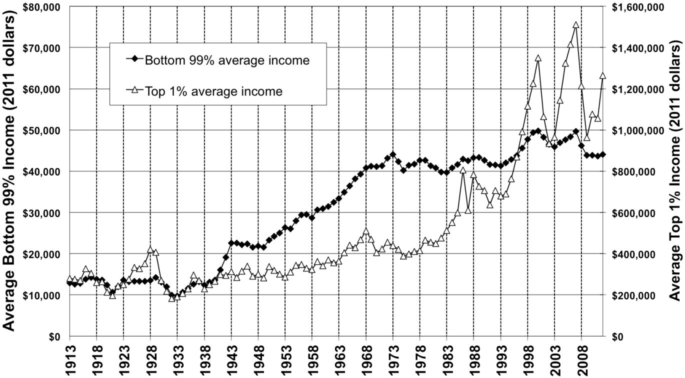

Source: Piketty and Saez, 2003 updated to 2012. Series based on pre-tax cash market income including realized capital gains and excluding government transfers. 2012 data based on preliminary statistics.

# The income distribution over time (4)




Source: Piketty and Saez, 2003 updated to 2011. Series based on pre-tax cash market income including or excluding realized capital gains, and always excluding government transfers.

# The income distribution over time (5)



# The income distribution over time (6)

Auten and Splinter (2017)

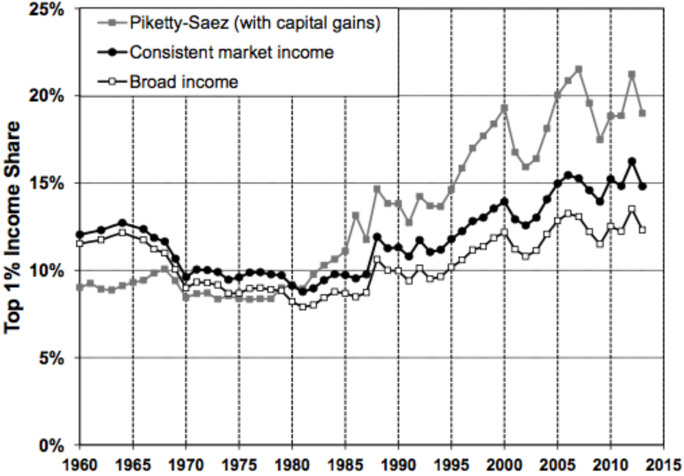
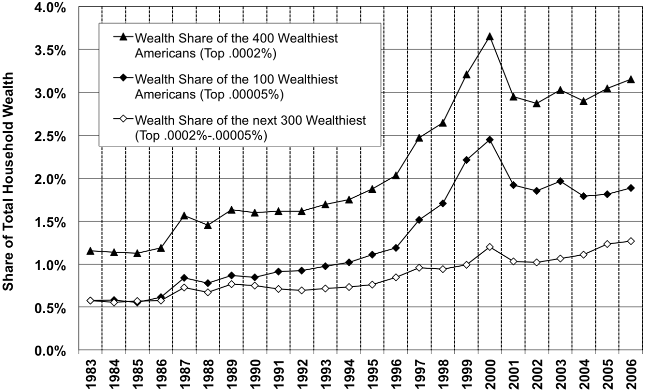



**Figure 3: Comparison of top 1% income shares**

*Notes*: Piketty and Saez series includes capital gains, where top one percent thresholds are defined by income excluding capital gains. Broad income is consistent market income plus government transfers. Adjustments used to estimate consistent market income and broad income are listed in Tables 1 and A1 and described in detail in the online appendix. All measures are pre-tax.

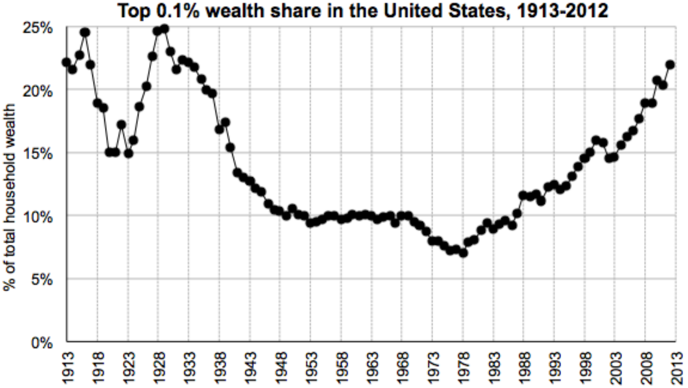
*Sources:* Authors' calculations, IRS, BEA, and Piketty and Saez (2003 and updates).

# The distribution of wealth over time (1)



# The distribution of wealth over time (2)

Saez and Zucman (2016)

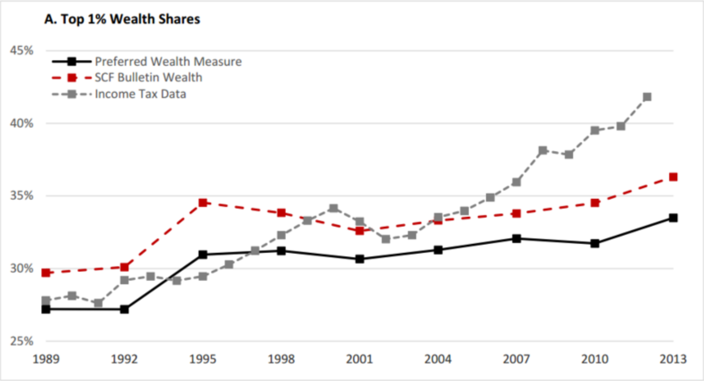



This figure depicts the share of total household wealth held by the 0.1% richest families, as estimated by capitalizing income tax returns. In 2012, the top 0.1% includes about 160,000 families with net wealth above $20.6 million. Source: Appendix Table B1.

# The distribution of wealth over time (3)

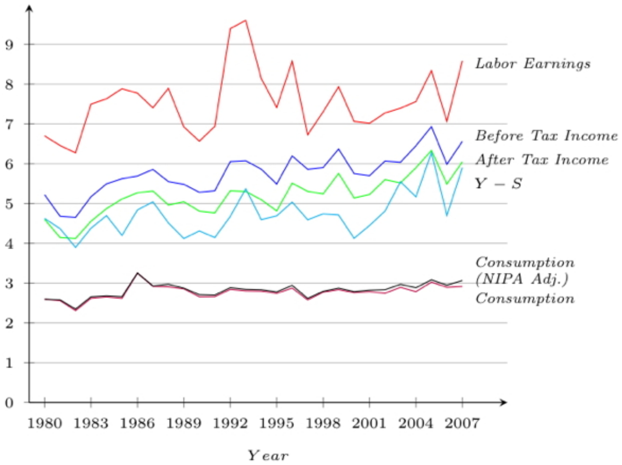
Bricker, Henriques, Krimmel, and Sabelhaus (2017)

**Figure 5. Top Wealth Shares, 1989-2013**



# Consumption inequality (1)

**Figure 1:** Trends in Inequality



# Consumption inequality (2)

- You want income in order to consume - maybe what is important is consumption inequality
- Much less increase here (but depends **a lot** on which data you look at)
- How could the two series (income and consumption inequality) not have the same trend?
    - The Permanent Income Hypothesis suggests that if increases in inequality are because of transitory effects, consumption inequality won't be affected
    - Maybe income measures are missing something - e.g. fringe benefits, which have become **much** more important over time

# Why is inequality increasing?

- Skill-biased technological change
  - e.g., as evidenced by the rising college wage premium
  - Instagram, WhatsApp
- Superstar effects
  - e.g., Barry Bonds
- Assortative mating
- Labor's share of income falling
- Increases in income mobility
- Increases in income shocks
- Changing demographics (aging population, immigration)
- But really, **we don't know**

# Transitory vs. persistent inequality

JASON DEBACKER
*Middle Tennessee State University*

BRADLEY HEIM
*Indiana University*

VASIA PANOUSI
*Board of Governors of the Federal Reserve System*

SHANTHI RAMNATH
*U.S. Department of the Treasury*

IVAN VIDANGOS
*Board of Governors of the Federal Reserve System*

## *Rising Inequality: Transitory or Persistent? New Evidence from a Panel of U.S. Tax Returns*

**ABSTRACT** We use a new, large, and confidential panel of tax returns to study the persistent-versus-transitory nature of rising inequality in male labor earnings and in total household income, both before and after taxes, in the United States over the period 1987–2009. We apply various statistical decomposition methods that allow for different ways of characterizing persistent and transitory income components. For male labor earnings, we find that the entire increase in cross-sectional inequality over our sample period was driven by an increase in the dispersion of the persistent component of earnings. For total household income, we find that most of the increase in inequality reflects an increase in the dispersion of the persistent income component, but the transitory component also appears to have played some role. We also show that the tax system partly mitigated the increase in income inequality, but not sufficiently to alter its broadly increasing trend over the period.

# Transitory vs. persistent inequality

Goal of this paper:

- Decompose the variance in income into its component sources
- Particular focus on the variance not explained by business cycles, age, or other observables
- Show how this has changed over time
- Do this in a way that incorporates the upper tail of the income distribution
- Show the effects of the tax system on income inequality

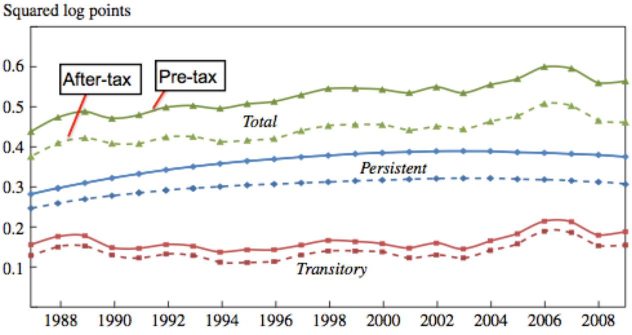
# Transitory vs. persistent inequality

What to we mean by transitory versus persistent?

- Transitory shocks to income are those that affect your income for only a short period of time
  - e.g., A one-time bonus
- Persistent shocks are those that last for a longer period of time (i.e., several years)
  - Some could be permanent
    - e.g., a step up in pay scale due to an advanced degree
  - Some are just persistent
    - e.g., an increase in demand for a particular skill over several years

# Transitory vs. persistent inequality

**Figure 9.** ECM Decomposition of Cross-Sectional Variance in Pre-Tax and After-Tax Household Income, All Households, 1987–2009




Source: Authors' calculations using SOI data.

# Transitory vs. persistent inequality

Why important to decompose?

- May say something about cause of inequality
  - E.g., if permanent component driving increase, then not increases in mobility
- Says something about welfare effects
  - E.g., if transitory shocks driving increase, these can be insured against by saving/borrowing

# Transitory vs. persistent inequality

DHPRV (2013) use a number of methods to decompose the variance in earnings:

- Non-parametric methods
  - Volatility (i.e., standard deviation in percentage changes) in income over different time horizons
  - Variance decomposition
    - Find variance in income over some time horizon (e.g. 5 years) - call this the "persistent" variance.
    - Find variance of (income less mean income over the time horizon) - call this the "transitory" variance
- Parametric Methods
  - Error components models

# Model of Income

Let $y^i_{a,t}$ = log income of individual $i$ at age $a$ in year $t$
Income is determined as:

$$y^i_{a,t} = g(\zeta; X^i_{a,t}) + \xi^i_{a,t}, \tag{1}$$

where:

- $\zeta$ is a vector of (possibly year dependent) parameters
- $X^i_{a,t}$ is a vector of observed characteristics
- $g(\cdot)$ is the component of log income that is common to all individuals conditional on $X^i_{a,t}$
- $\xi^i_{a,t}$ is the unobserved error term

# Error-Components Model

ECMs put structure on this error term to estimate the process underlying this unobserved component of income.
Assume (for the stationary model of residual earnings):

$$\xi^i_{a,t} = \alpha^i + p^i_{a,t} + \tau^i_{a,t} \tag{2}$$

- $\alpha^i$ is the permanent component (i.e., an unobserved fixed effect)
- $p^i_{a,t}$ is the persistent component
- $\tau^i_{a,t}$ is the transitory component

# ECM

Some further structure:

- The persistent component follows an AR(1) process:

$$p^i_{a,t} = \psi p^i_{a-1,t-1} + \eta^i_{a,t} \tag{3}$$

- The transitory component follows an MA(2) process:

$$\tau^i_{a,t} = \varepsilon^i_{a,t} + \theta_1 \varepsilon^i_{a-1,t-1} + \theta_2 \varepsilon^i_{a-2,t-2} \tag{4}$$

- Shocks are distributed as:

$$\alpha^i \sim \text{i.i.d.}(0, \sigma^2_\alpha), \eta^i_{a,t} \sim \text{i.i.d.}(0, \sigma^2_\eta), \varepsilon^i_{a,t} \sim \text{i.i.d.}(0, \sigma^2_\varepsilon)$$

# ECM

Some notes:

- This is a lot of structure to put on the model
- It is somewhat flexible (e.g., it may turn out that the process is MA(1) and thus $\theta_2 = 0$)
- But do note that the parameters are not index by $t$, so assuming some stationarity (this is relaxed in DHPRV (2013))
- You want to look at data and have that help you inform the parametric specification (i.e., plot residuals - what to they look like? What does the auto-covariance structure look like? Are the relationships stable over time?)

# Identification

What would this model of the error term imply for the structure of the data?

1. $E(\xi^i_{a,t}) = E(\alpha_i + p^i_{a,t} + \tau^i_{a,t}) = 0$
2. $cov(\xi^i_{a,t}, \xi^i_{a+k,t+k}) =$
   $E(\xi^i_{a,t} * \xi^i_{a+k,t+k}) + \underbrace{E(\xi^i_{a,t})}_{=0}\underbrace{E(\xi^i_{a+k,t+k})}_{=0} = E(\xi^i_{a,t} * \xi^i_{a+k,t+k})$

# Identification

Further expansion of the $cov(\xi^i_{a,t}, \xi^i_{a+k,t+k})$:

$$
\begin{aligned}
cov(\xi^i_{a,t}, \xi^i_{a+k,t+k}) &= E(\xi^i_{a,t} * \xi^i_{a+k,t+k}) \\
&= E\left[(\alpha_i + p^i_{a,t} + \tau^i_{a,t})(\alpha_i + p^i_{a+k,t+k} + \tau^i_{a+k,t+k})\right] \\
&= E\Big[\alpha^i\alpha^i + \alpha^i p^i_{a,t} + \alpha^i \tau^i_{a,t} + \alpha^i p^i_{a+k,t+k} + p^i_{a,t}p^i_{a+k,t+k} + \\
&\qquad \tau^i_{a,t}p^i_{a+k,t+k} + \alpha^i\tau^i_{a+k,t+k} + p^i_{a,t}\tau^i_{a+k,t+k} + \tau^i_{a,k}\tau^i_{a+k,t+k}\Big] \\
&= \underbrace{E(\alpha^i\alpha^i)}_{=var(\alpha^i)} + \underbrace{E(\alpha^i p^i_{a,t})}_{=0} + \underbrace{E(\alpha^i\tau^i_{a,t})}_{=0} + \underbrace{E(\alpha^i p^i_{a+k,t+k})}_{=0} + \\
&\qquad \underbrace{E(p^i_{a,t}p^i_{a+k,t+k})}_{=cov(p^i_{a,k}p^i_{a+k,t+k})} + \underbrace{E(\tau^i_{a,t}p^i_{a+k,t+k})}_{=0} + \underbrace{E(\alpha^i\tau^i_{a+k,t+k})}_{=0} + \\
&\qquad \underbrace{E(p^i_{a,t}\tau^i_{a+k,t+k})}_{=0} + \underbrace{E(\tau^i_{a,k}\tau^i_{a+k,t+k})}_{=cov(\tau^i_{a,k},\tau^i_{a+k,t+k})}
\end{aligned}
$$

# Identification

$$cov(\xi^i_{a,t},\xi^i_{a+k,t+k}) = var(\alpha^i) + \underbrace{cov(p^i_{a,k}p^i_{a+k,t+k})}_{\rho^k var(p^i_{a,t})} + cov(\tau^i_{a,k}\tau^i_{a+k,t+k})$$
$$= \sigma^2_\alpha + \rho^k var(p^i_{a,t}) + cov(\tau^i_{a,k},\tau^i_{a+k,t+k})$$

where,

$$cov(\tau^i_{a,k}\tau^i_{a+k,t+k}) = \begin{cases} \sigma^2_\varepsilon, & \text{if } k=0, a=1 \\ (1+\theta_1^2)\sigma^2_\varepsilon, & \text{if } k=0, a=2 \\ (1+\theta_1^2+\theta_2^2)\sigma^2_\varepsilon, & \text{if } k=0, a\geq 3 \\ \theta_1\sigma^2_\varepsilon, & \text{if } k=1, a=1 \\ (\theta_1+\theta_1\theta_2)\sigma^2_\varepsilon, & \text{if } k=1, a\geq 2 \\ \theta_2\sigma^2_\varepsilon, & \text{if } k=2 \\ 0, & \text{if } k>2 \end{cases}$$

and

$$var(p^i_{a,t}) = \begin{cases} \sigma^2_\eta \frac{1-\psi^{2a}}{1-\psi^2} & \text{if } a\geq 2, t = \text{First year in data} \\ \sigma^2_\eta, & \text{if } a=1 \\ \psi^2 var(p^i_{a-1,t\_1}) + \sigma^2_\eta & \text{if } a\geq 2, t \neq \text{First year in data} \end{cases}$$

# Identification

Theoretical covariances:

$$
\begin{aligned}
cov(a,t,k;\Theta) = \sigma_\alpha^2 + \psi^k var(p^i_{a,t}) + & \\
& \mathbb{1}[k=0](1+\mathbb{1}[a\geq 2]\theta_1^2 + \mathbb{1}[a\geq 3]\theta_2^2)\sigma_\varepsilon^2 + \\
& \mathbb{1}[k=1](\theta_1 + \mathbb{1}[a\geq 2]\theta_1\theta_2)\sigma_\varepsilon^2 + \\
& \mathbb{1}[k=2]\theta_2\sigma_\varepsilon^2
\end{aligned}
$$

- Analytical solution for the covariances that we observe in the data
- $\implies$ moment conditions (for each $a$, $t$, $k$):

$$
g_{a,t,k}(\Theta_0) = E\left[\underbrace{cov(\xi^i_{a,t},\xi^i_{a+k,t+k})}_{\text{covariances from data}}\right] - cov(a,t,k;\Theta_0) = 0
$$

# GMM Estimator

We will use the sample analogue of the theoretical moment conditions:

$$\tilde{g}_{a,t,k}(\Theta_0) = \frac{\sum_{i=1}^{n}(\xi_{a,t}^{i}\xi_{a+k,t+k}^{i})}{n-1} - cov(a,t,k;\Theta_0) = 0$$

Define $g(\Theta)$ as the vector containing all the $g_{a,t,k}(\Theta)$ moment conditions.
The GMM estimator is then:

$$\hat{\Theta} = arg\min_{\Theta} g(\Theta)^T W g(\Theta),$$

where $W$ is a weighting matrix