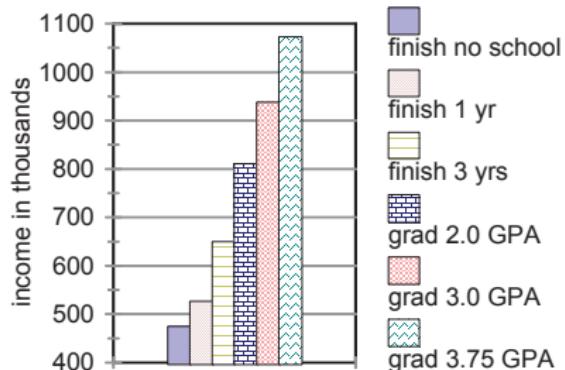


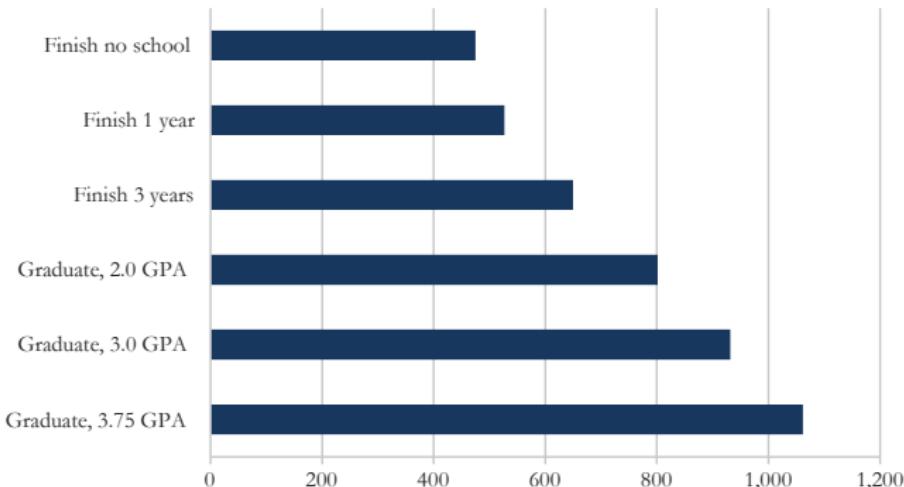
**Figure 2 Discounted Expected Lifetime  
Earnings,  $VN(t')$**



Source: Stinebrickner and Stinebrickner (2013).

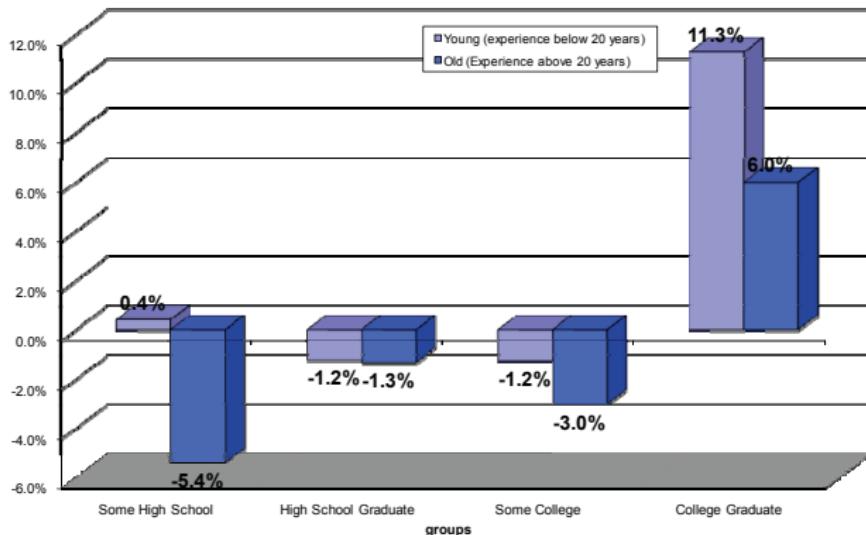
### **Discounted Expected Lifetime Earnings, $VN(t')$**

(Income in thousands)



Source: Author's calculations using numbers inferred from text in Stinebrickner and Stinebrickner (2013).

## Change in real weekly wages of US-born workers by group, 1990-2006

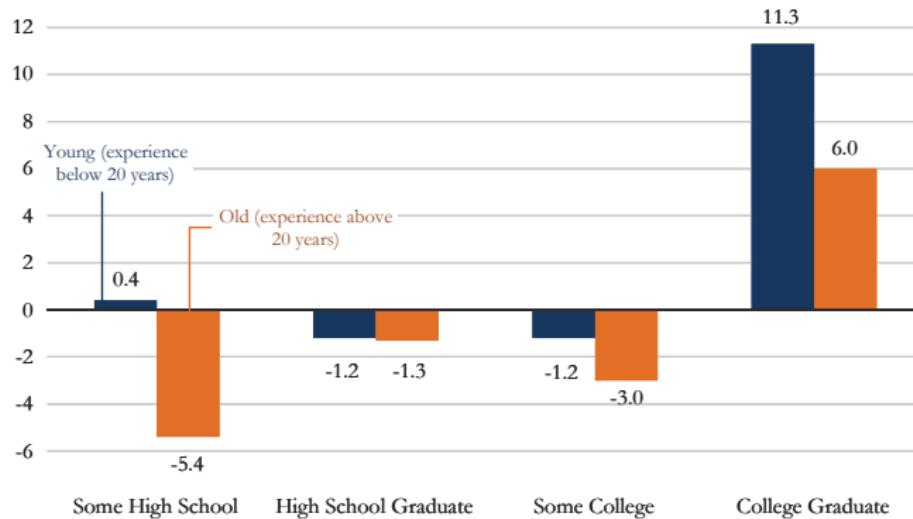


Source: Ottaviano and Peri (2008).

## Flattening a 3D Chart

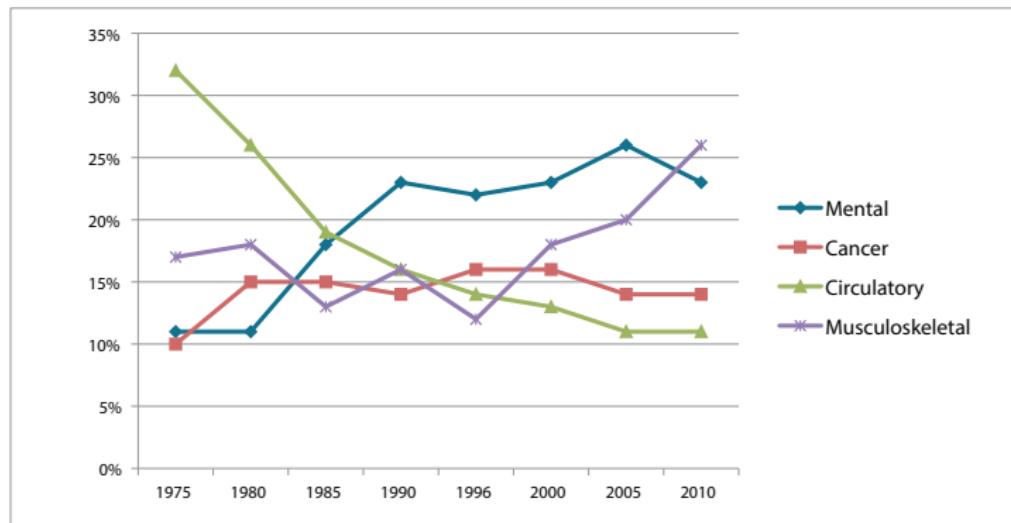
Change in real weekly wages of US-born workers by group, 1990–2006

(Percent)



*Figure 6A*  
**A Spaghetti Chart**

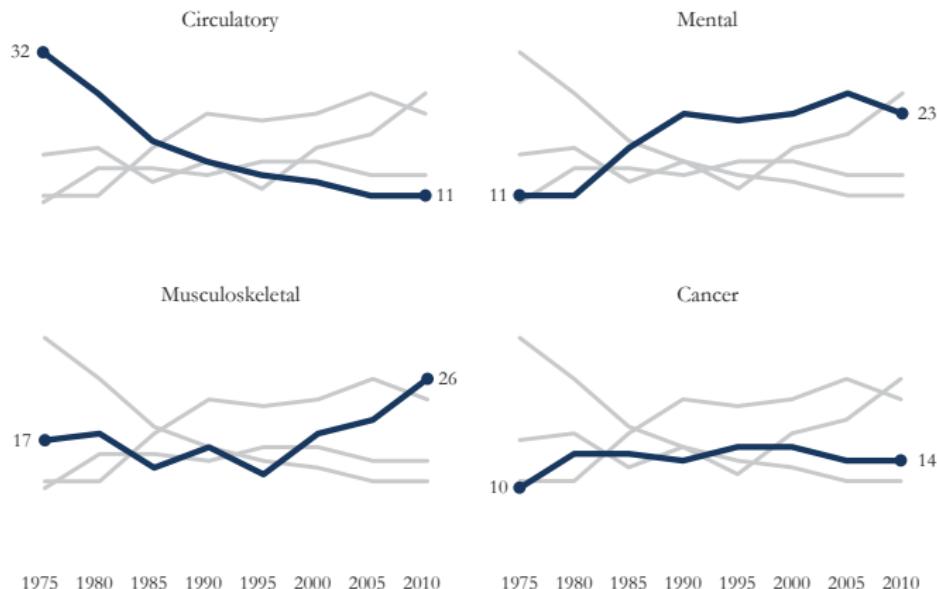
27. Initial DI Worker Awards by Major Cause of Disability—Calendar Years 1975-2010



Source: Social Security Advisory Board (2012).

*Figure 6B*  
Revising the Spaghetti Chart

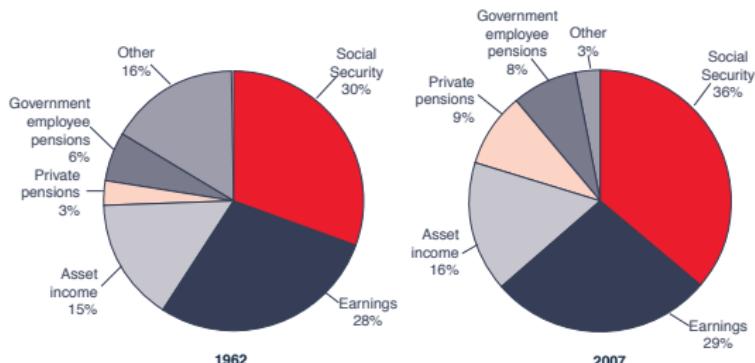
Initial DI Worker Awards by Major Cause of Disability—  
Calendar Years 1975–2010  
(Percent)



*Figure 9A*  
**Two Pie Charts for Comparison**

### Shares of Aggregate Income, 1962 and 2007

Aggregate income, by source

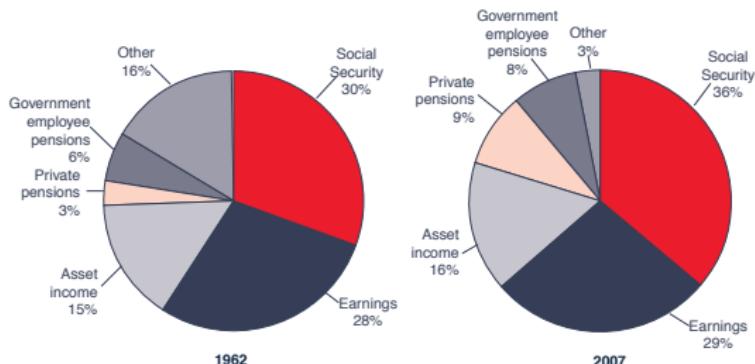


Source: Social Security Administration (2009).

*Figure 9A*  
**Two Pie Charts for Comparison**

### Shares of Aggregate Income, 1962 and 2007

Aggregate income, by source



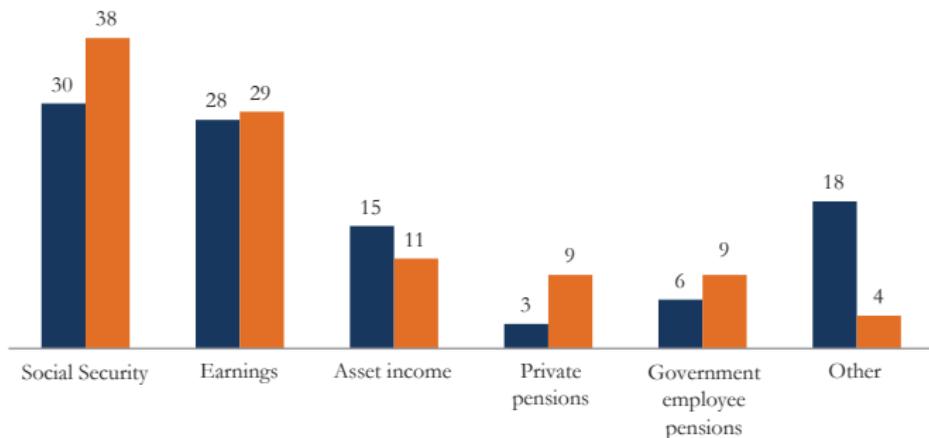
Source: Social Security Administration (2009).

*Figure 9B*

### Alternative to a Pie Chart: A Paired Column Chart

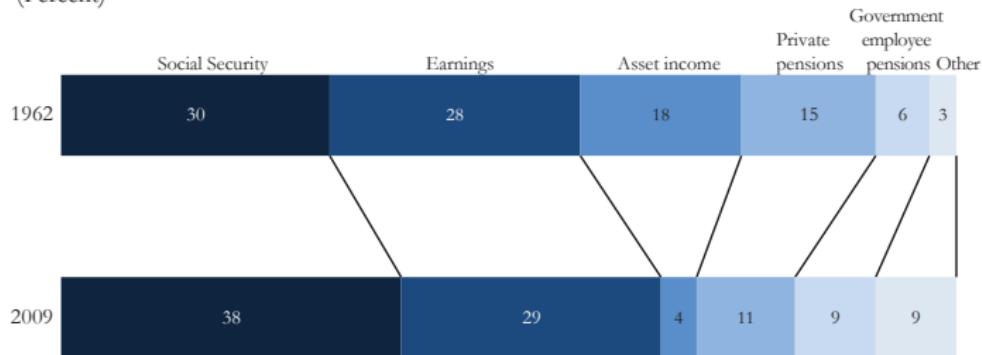
Shares of Aggregate Income, 1962 and 2009  
(Percent)

■ 1962 ■ 2009

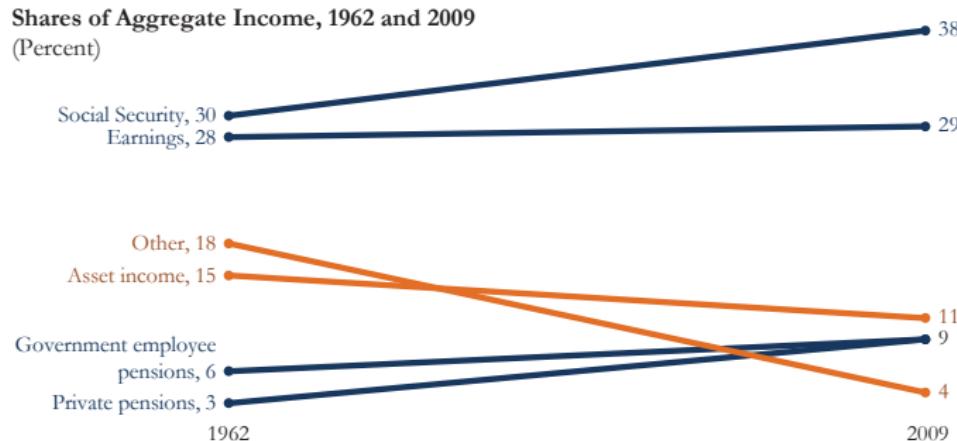


*Figure 9C*  
Alternative to a Pie Chart: A Stacked Bar Chart

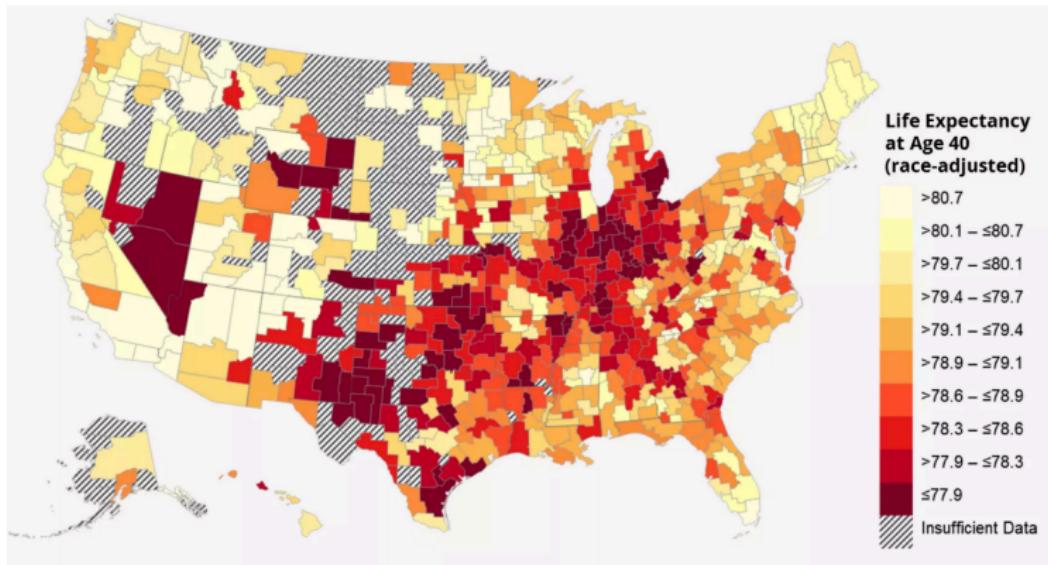
Shares of Aggregate Income, 1962 and 2009  
(Percent)



*Figure 9D*  
Alternative to a Pie Chart: The Slope Chart

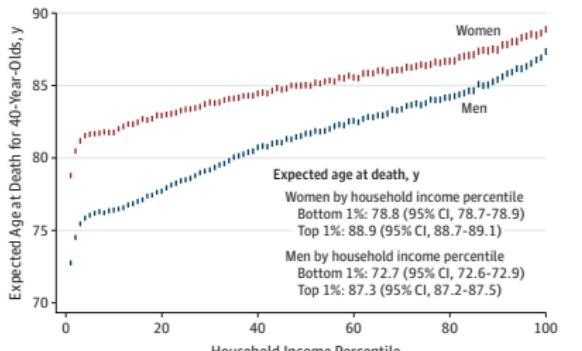


# A data set in pictures: Chetty et al, JAMA, 2016



# A data set in pictures: Chetty et al, JAMA, 2016

Figure 2. Race- and Ethnicity-Adjusted Life Expectancy for 40-Year-Olds by Household Income Percentile, 2001-2014



Mean household income  
in thousands, \$<sup>a</sup>

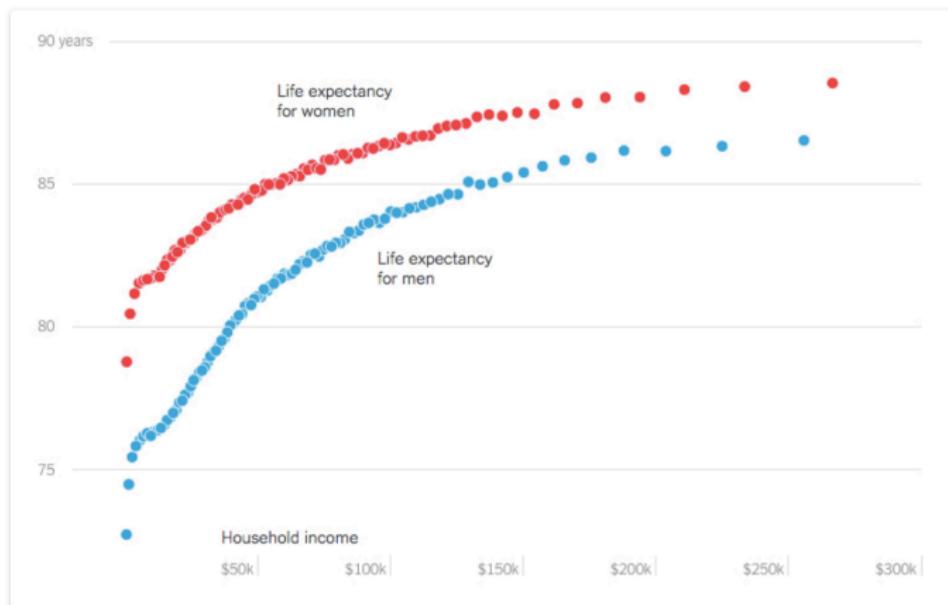
Women	24	45	71	112	1.9 million
Men	26	50	77	119	2.0 million

Life expectancies were calculated using survival curves analogous to those in Figure 1. The vertical height of each bar depicts the 95% confidence interval. The difference between expected age at death in the top and bottom income percentiles is 10.1 years (95% CI, 9.9-10.3 years) for women and 14.6 years (95% CI, 14.4-14.8 years) for men. To control for differences in life expectancies across racial and ethnic groups, race and ethnicity adjustments were calculated

using data from the National Longitudinal Mortality Survey and estimates were reweighted so that each income percentile bin has the same fraction of black, Hispanic, and Asian adults.

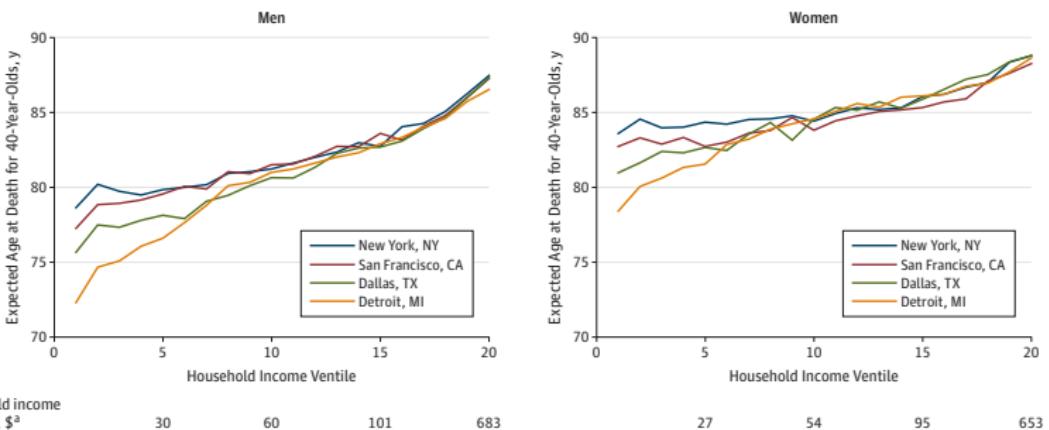
<sup>a</sup> Averaged across years and ages. The data are in thousands unless otherwise indicated.

# A data set in pictures: Chetty et al, JAMA, 2016, as rendered in NYT



# A data set in pictures: Chetty et al, JAMA, 2016

Figure 4. Race- and Ethnicity-Adjusted Life Expectancy by Income Ventile in Selected Commuting Zones, 2001-2014



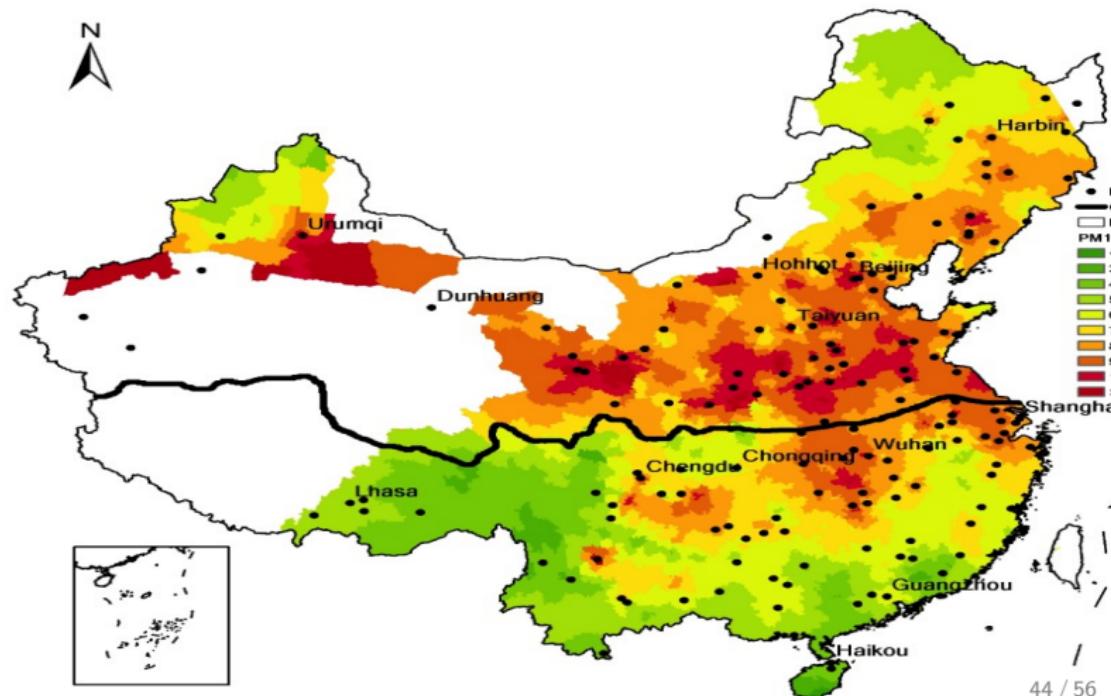
Estimates of race- and ethnicity-adjusted expected age at death for 40-year-olds computed by income ventile (5 percentile point bins).

<sup>a</sup> Averaged across years and ages.

# Visualizing a research strategy

**Figure 1**

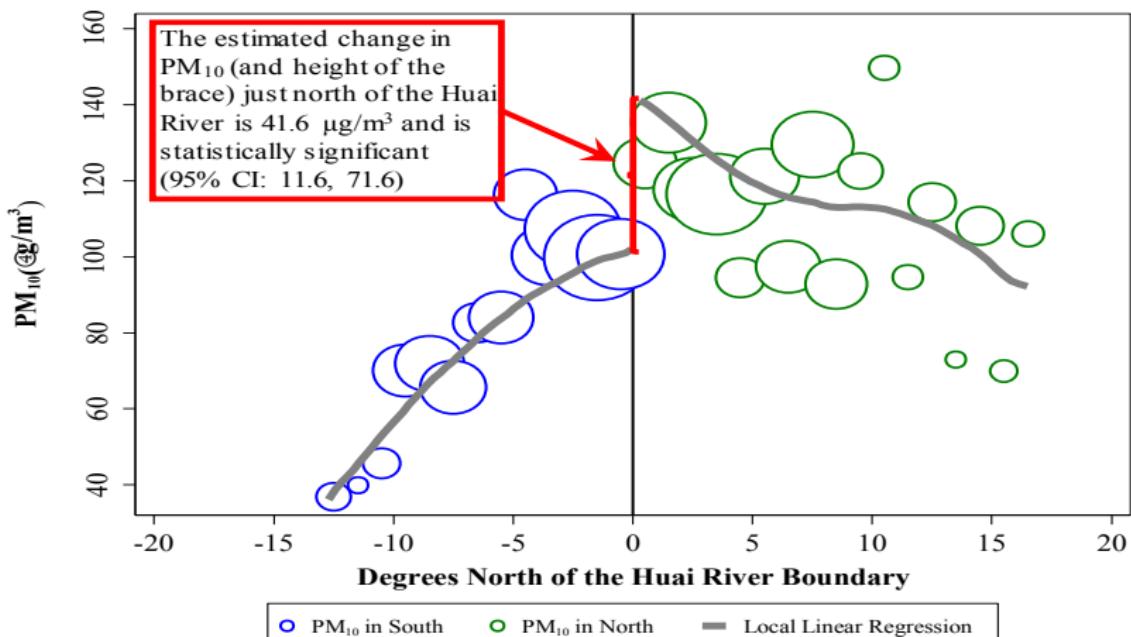
Pollution in China and the Huai River/Qinling Mountain Range



# Visualizing a research strategy

**Figure 2**

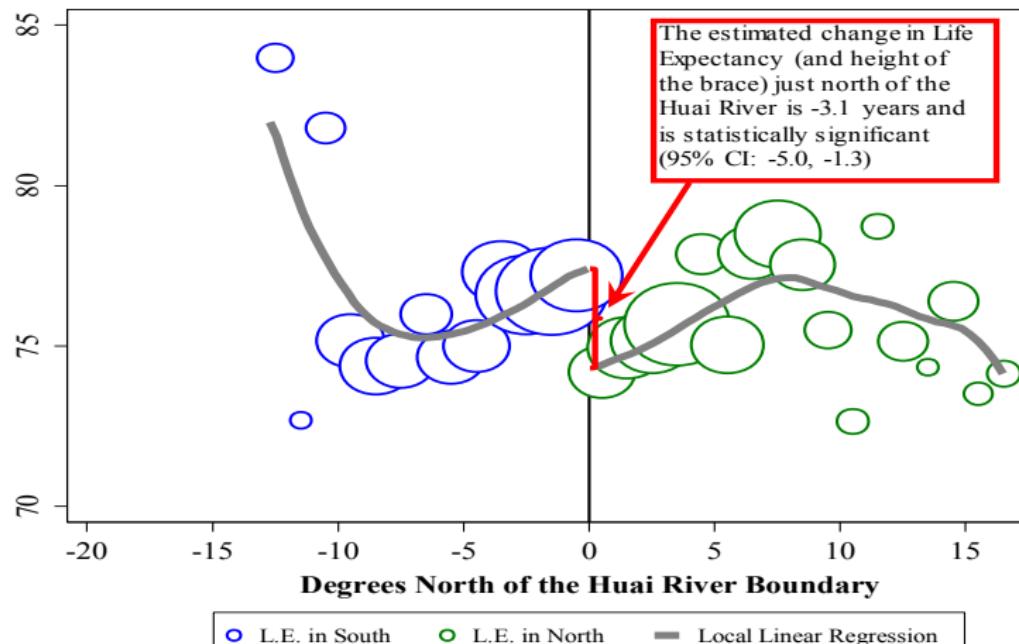
Particulate Matter Levels ( $\text{PM}_{10}$ ) South and North of the Huai River Boundary



# Visualizing a research strategy

**Figure 3**

Life Expectancy South and North of the Huai River Boundary



## Tables

- Use the same principle:
  - Show the data
  - Don't lie about it
  - Focus
- Which translates in:
  - Report the important numbers, not all the coefficients
  - Keep the number of significant digits down
  - No vertical lines,
  - very few horizontal lines: don't box results (3 lines are usually enough)
  - in doubt, align left
  - Report the statistics that the reader will need, not necessary those that come by default.
  - Sample questions to ask yourself:
    - Do you need the R<sup>2</sup>?
    - Are there important tests to report (say between equality of two coefficients)?
    - You may need the mean in the control group?

Table 7: Impact of the reform on fake beneficiaries: Evidence from matching of nrega.nic.in job cards with SECC census

	All job cards (as of April 2014)	Job cards with at least one working member		
		Intervention period July 2012-March		Post intervention Apr 2013 - March
		2013 (1)	2014 (2)	2014 (3)
<b>Panel A: Match Rate for job cards with one member only</b>				
Treatment	0.0187** (0.00741)	0.0181** (0.00766)	0.0107 (0.00696)	
Observations	3,095	2,868	2,922	
Mean in Control	0.644	0.673	0.698	
<b>Panel B: Match Rate for job cards with two members or more</b>				
Treatment	0.0135** (0.00613)	0.0126 (0.00764)	0.0104 (0.00732)	
Observations	3,093	2,836	2,906	
Mean in Control	0.243	0.282	0.286	

Note: The unit of observation is a GP. The dependent variable is the fraction of job cards from nrega.nic.in matched by name with households from the SECC census. A job card with two members or more is matched when at least two members have been matched by name with a census household. The nrega.nic.in data was extracted from the nrega.nic.in server, it covers the period from July 2011 to March 2014. Treatment is a dummy which is equal to one for the blocks selected for the intervention. All specifications include district fixed effects.

## Stargazer

- In practice, in R, use the "stargazer" package
- It will create standardized tables, can output them for you in html, latex, word, ASCII
- The default is not bad
- But you can customize it to chose what statistics to include, chose what coefficients to include, etc.
- <https://cran.r-project.org/web/packages/stargazer/vignettes/stargazer.pdf>
- <http://jakeruss.com/cheatsheets/stargazer.html>

# Default output

TABLE 1. Results

	<i>Dependent variable:</i>	
	ftvoteshare	
	(1)	(2)
fncandidates	0.134*** (0.007)	
RESprior		0.059** (0.024)
Constant	0.012 (0.010)	0.094*** (0.019)
Observations	372	372
R <sup>2</sup>	0.532	0.016
Adjusted R <sup>2</sup>	0.530	0.013
Residual Std. Error (df = 370)	0.153	0.222
F Statistic (df = 1; 370)	420.093***	6.056**

*Note:* \*p<0.1; \*\*p<0.05; \*\*\*p<0.01

## Improve on it

- Transparent label names for dependent and independent variables
- Choose the coefficients you would like to include
- No need to include both R squared and adjusted R squared
- Maybe the need of the dependent variable in the control group, rather than the constant?

# Visualizing regression results

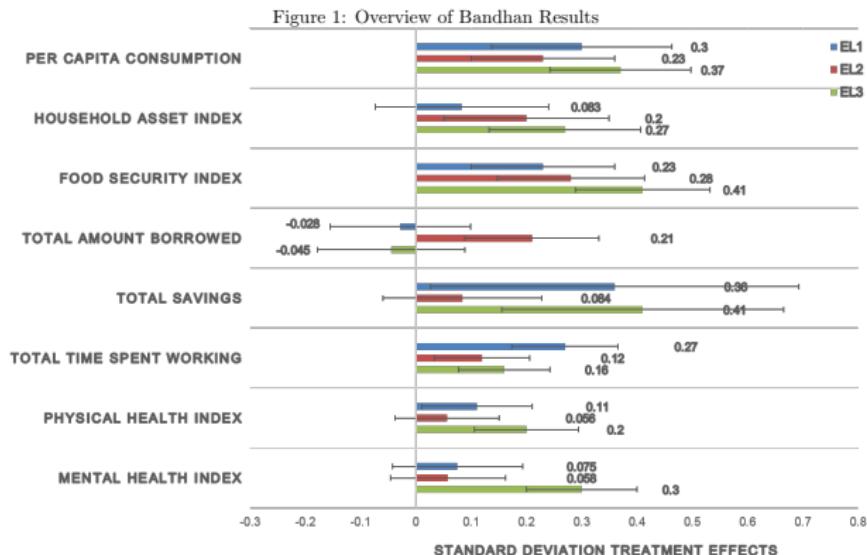
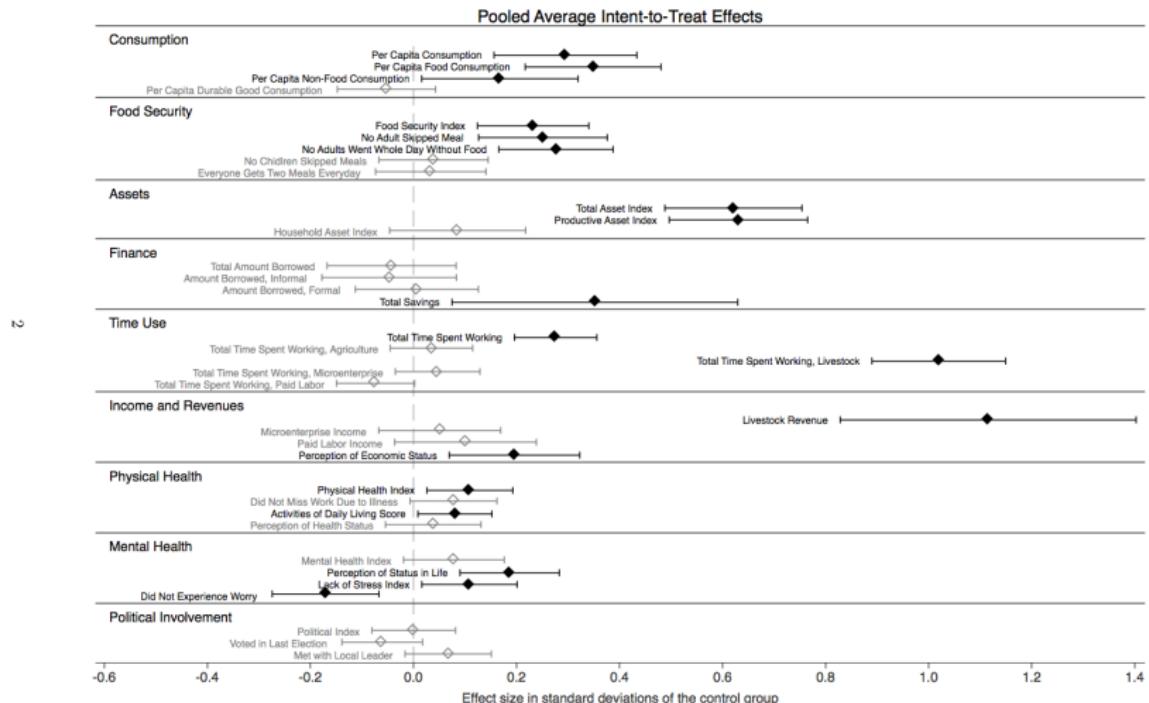


Figure 2: Endline 1-At a Glance



This figure summarizes the treatment effects. Here, treatment effects on continuous variables are presented in standard deviation units.  
Each line shows the OLS point estimate and 95% confidence interval for that outcome.

## Tools and resources

- R is a great visualization tool (especially ggplot; Need to pick up a good book to chose colors, etc.)
- Yau: Flowing data has many tutorials for how to do things in R (<http://flowingdata.com/category/tutorials/>) and a 4 weeks mini-course you can take at your own pace.
- R handles/produces maps as well.
- At this site you can find a Tufte's charts in R  
<http://motioninsocial.com/tufte/>

## References

- Edward Tufte *Visual Display of Quantitative Information* Graphics Press, 2013 (second edition)
- Jonathan Schwabish “An Economist’s guide to visualizing data” *Journal of Economic Perspective*, 2014, vol 28, number 1
- Banerjee, Chattopadhyay, Duflo, Shapiro “The long term impact of a graduation program: Evidence from West Bengal” Mimeo, MIT
- Duflo, Esther “Schooling and Labor Market Consequences of School construction in Indonesia” American Economic Review, 2001
- Chetty, Raj, Michael Stepner, Sarah Abraham, Shelby Lin, Benjamin Scuderi, Nicholas Turner, Augustin Bergeron, and David Cutler. “The association between income and life expectancy in the United States, 2001-2014.” JAMA 315, no. 16 (2016): 1750-1766.

## References

- Bui, Quoctrung and Neil Irwin. "The Rich Live Longer Everywhere. For the Poor, Geography Matters. New York Times April 11, 2016.
- Chen, Yuyu, Avraham Ebenstein, Michael Greenstone, and Hongbin Li. "Evidence on the Impact of Sustained Exposure to Air Pollution on Life Expectancy from China's Huai River Policy."
- Lee, David S., Enrico Moretti, and Matthew J. Butler. "Do voters affect or elect policies? Evidence from the US House." The Quarterly Journal of Economics (2004): 807-859.
- Klerman, Jacob Alex, and Caroline Danielson. "The Transformation of the Supplemental Nutrition Assistance Program. Journal of Policy Analysis and Management, vol. 30, no. 4, 2011, pp. 863888

## References

- Stinebrickner, Ralph, and Todd Stinebrickner. "Academic Performance and College Dropout: Using Longitudinal Expectations Data to Estimate a Learning Model. Western University CIBC Working Paper 2013-5 ( July),  
[http://economics.uwo.ca/cibc/workingpapers\\_docs/wp2013/Stinebrickner\\_Stinebrickner05.pdf](http://economics.uwo.ca/cibc/workingpapers_docs/wp2013/Stinebrickner_Stinebrickner05.pdf). 2013.
- Ottaviano, Gianmarco I. P., and Giovanni Peri. Immigration and National Wages: Clarifying the Theory and the Empirics. NBER Working Paper 14188 ( July). 2008
- Social Security Advisory Board. Aspects of Disability Decision Making: Data and Materials. Social Security Administration (February), [http://www.ssab.gov/Publications/Disability/GPO\\_Chartbook\\_FINAL\\_06122012.pdf](http://www.ssab.gov/Publications/Disability/GPO_Chartbook_FINAL_06122012.pdf). 2012

## References

- Social Security Administration. Fast Facts Figures About Social Security, 2009. Social Security Administration ( July). [http://www.ssa.gov/policy/docs/chartbooks/fast\\_facts/2009/fast\\_facts09.pdf](http://www.ssa.gov/policy/docs/chartbooks/fast_facts/2009/fast_facts09.pdf). 2009
- Klerman, Jacob Alex, and Caroline Danielson. "The Transformation of the Supplemental Nutrition Assistance Program. Journal of Policy Analysis and Management 30(4): 863 88. 2011
- Hanson, Gordon H. 2012. The Rise of Middle Kingdoms: Emerging Economies in Global Trade. Journal of Economic Perspectives 26(2): 41 64.