Paper Examples

- Katz and Murphy, QJE (1992)

- Goldin and Margo, The Great Compression, QJE (1992)

- Goldin and Katz, The Power of the Pill, JPE (2002)

- Costa and Kahn, Cowards and Heroes, QJE (2003)

Tables

- Use the notes: each table/figure should be self-contained
  - Remember that most readers will only skim the paper
  - BUT, don't use the notes as a substitute for a data appendix
- Give descriptive labels for the variables
- Be clear what the dependent variable is
- If it's Eq. 2 in the paper, then say so
- Don't present coefficients that no one is interested in. If you're not going to discuss the coefficients, don't waste space

TABLE VI
DECOMPOSITION OF THE CHANGE IN THE WAGE STRUCTURE, 1940 TO 1950

Using 1950 as the base year

Difference in percentiles	Total change log (wages)[a]	Due to change in		
		Observed quantities	Observed prices	Distribution of residuals
90-10	−0.354	−0.172	−0.167	−0.016
90-50	−0.202	−0.088	−0.077	−0.036
50-10	−0.153	−0.083	−0.090	+0.020

Using 1940 as the base year

Difference in percentiles	Total change log (wages)[a]	Due to change in		
		Observed quantities	Observed prices	Distribution of residuals
90-10	−0.354	+0.007	−0.143	−0.219
90-50	−0.202	−0.014	−0.074	−0.114
50-10	−0.153	+0.021	−0.069	−0.105

a. Total change for the 90-10 row, for example, is the (log wage at the ninetieth percentile − log wage at the tenth percentile)$_{1950}$ − (log wage at the ninetieth percentile − log wage at the tenth percentile)$_{1940}$.

Notes. The decomposition is described in the text. The wage equation (1) was estimated for 1940 and 1950 on the sample of white male wage and salary earners, 18 to 64 years old, whose (full-time) weekly wage and salary earnings were more than one half the current minimum wage. The variables included are experience (= age − years of school − 6), experience squared, years of school, schooling squared, an interaction of schooling and experience, dummy variables for high school and college graduation, for region, for urban residence, for marital status, for household headship, and for foreign birth.

Sources. 1940, 1950, and 1960 PUMS. See text for a description of the decomposition procedure.

Figures

- A figure is worth a thousand words
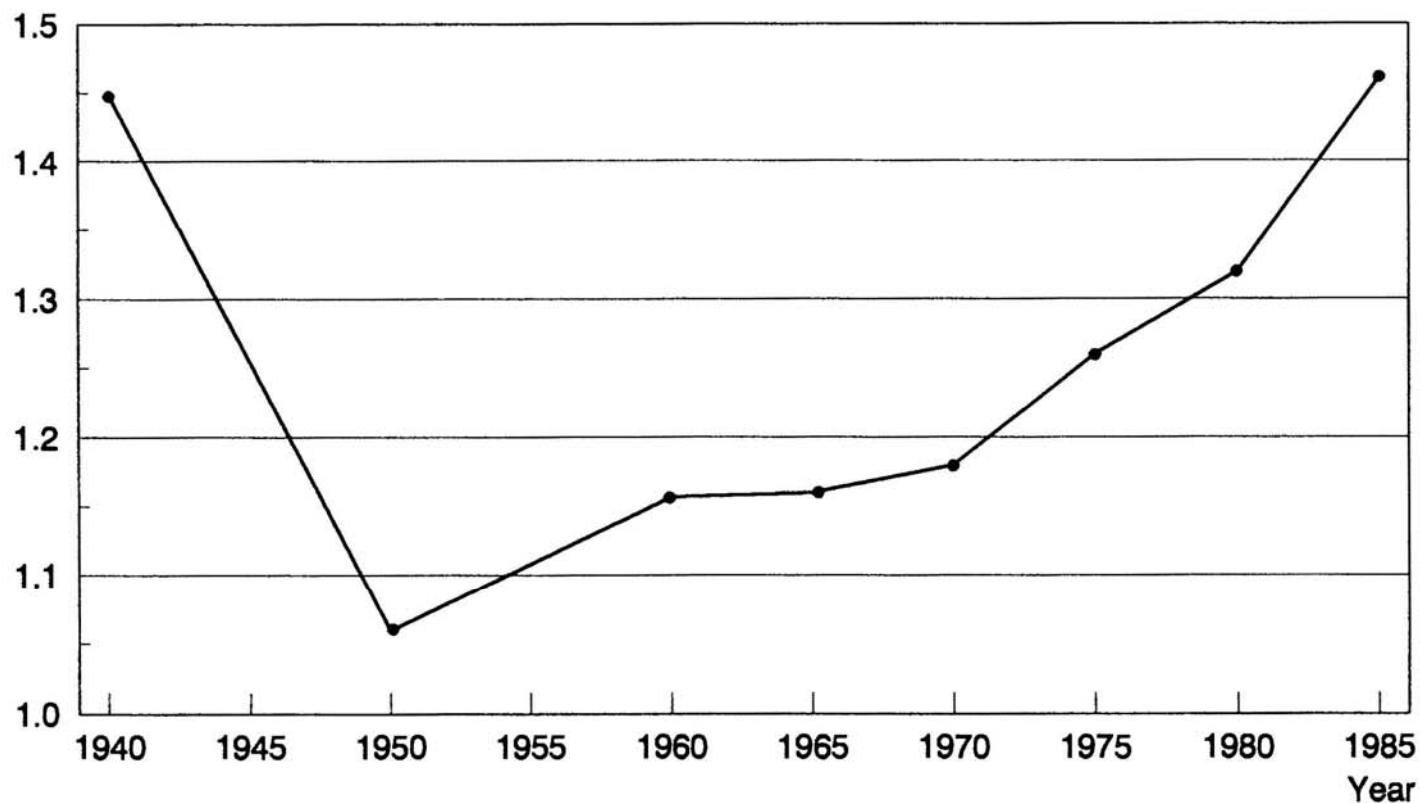- But sometimes you just can't use a figure …

Log(Wage), 90th-10th

FIGURE I

Wage Dispersion across the Past Half Century: Difference in the Log Wage at the Ninetieth and Tenth Percentiles, 1940 to 1985

Source. Table I.

Results Section

- Describe the main results in your tables
- Common errors:
  - Stick to the description; don't digress into comparisons with the literature
  - Don't digress into robustness results until you're done with the main results
  - Some material should be put in footnotes
- Don't put in everything you did but enough to indicate you were thorough

Data

- What data set using?
- What basic properties?
- How restrict sample?
- What variables constructed?
  - If it's complicated, full details can and should be given in a Data Appendix (Costa and Kahn 2003)
- Is there enough variation in data?
- Start paragraphs with most important point first
  - Details belong in the rest of the paragraph
  - Be to the point

II. The Data

The data we use in this paper come from a series of 25 consecutive March Current Population Surveys (CPSs) for survey years 1964 to 1988. These CPS data are from the March Annual Demographic Supplement and provide information on earnings and weeks worked in the calendar year preceding the March survey. These surveys provide wage and employment information on approximately 1.4 million workers for the 1963 to 1987 period. From these CPS data we create two samples: (1) a wage sample that we use to measure weekly wages of full-time workers by demographic group and (2) a count sample that we use to measure the amount of labor supplied by each of these demographic groups. The taxonomy we use divides the data into 320 distinct labor groups, distinguished by sex, education (less than 12, 12, 13–15, and 16 or more years of schooling), and 40 single-year potential experience categories (corresponding to the first 40 years since the estimated age of labor market entrance).[2]

The wage measure that we use throughout the paper is the average weekly wage of full-time workers (computed as total annual earnings divided by total weeks worked) within a gender-education-experience cell.[3] Our wage sample includes full-time wage and salary workers who participated in the labor force for at least 39 weeks in the calendar year prior to the March survey, worked at least one week, and did not work part year due to school, retirement, or military service. Self-employed workers and those working without pay were excluded from the wage sample. The sample includes individuals for whom the Census imputed wages but makes a correction for the fact that the imputation procedures changed between the 1975 and 1976 March CPS surveys.[4] Workers with top coded earnings were imputed annual earnings at 1.45 times the annual topcode amount. This correction is based on our estimates of the conditional average earnings of those with earnings above the topcode. In addition, we excluded workers with real weekly earnings below $67 in 1982 dollars (equal to one half of the 1982 real minimum wage based on a 40-hour week). As best as we can ascertain from experimentation, our results are not highly sensitive to these exclusion criteria.

The count sample includes all individuals who worked at least one week in the preceding year (regardless of whether they were wage and salary workers, self employed, or otherwise). We compute total hours worked for each cell in each year by computing the product of total annual hours (weeks worked times usual weekly hours) and the individual CPS sample weight for each individual in

Data Section (Cont)

- A place to put basic means if you can use them as something other than means (e.g. interesting in and of themselves or useful for talking about bias)
  - Another place to put means is in the results table
  - Another place to put means is a descriptive trends section
- Costa and Kahn (2003): "Table II illustrates the wide variation in shirking and mortality rates by states."

Where should I put the Data Section?

- Katz and Murphy (1992): Immediately after the introduction and before description of trends and supply and demand framework

- Costa and Kahn (2003): After Empirical Framework and after description of how Union Army worked

Trends or other Descriptive Stuff Section

- For some papers documenting a trend is particularly important

- For others need some basic background whether historical or other (e.g. how an industry works)

- e.g. Goldin and Katz (2002)
  - After intro: The Pill and Single Women
    - A: Diffusion
    - B: State Variation in Laws
    - C: Impact of State Laws
  - Frameworks to Understand comes after this section

Model/Framework Section

- If an empirical paper, should be related to empirical work
  - No point in presenting a theory you then don't test
  - Before you write this section up know what the empirical work is going to tell you
  - Although the way you think about the problem is going to be motivated by theory, some leads will be red herrings and there's no point inflicting these on the reader

Model/Framework (Cont.)

- Goldin and Katz (2002): "How could the diffusion of the pill have altered professional career investments by young women?"

  – theory shows that this is possible and then the next section presents empirical work to prove this

  – Note nice use of words to give intuition beforehand

timing was related to changes in state laws and a growing notion that young people could make their own decisions. But how could the diffusion of the pill have affected professional career investments by young women?

The diffusion of the pill among young, single women may have altered career decisions through two routes: *direct* and *indirect*. By the direct effect of the pill, we mean the reduction in the cost of marriage delay. The pill makes marriage delay and thus career investment cheaper, and women with greater "career ability" become more attractive marriage partners. By the indirect effect of the pill, we mean the lowering of the cost of a career through the marriage market. This effect, in contrast, operates through a thickening of the marriage market for those who delay marriage and leads to better matches for career women and some others. To simplify the discussion, we formally model the direct effect and give the intuition behind the indirect effect.[31]

Consider a cohort of n women and n men, each initially unmarried, in a two-period context with no discounting. Members of each sex agree on the ranking of the other in terms of marriage partners. Each man

Model/Framework (Cont.)

- Use enough words throughout so clear what each step in the model is

- Clearly state what the predictions are
  - Note that they could be ambiguous and what you're just estimating is the size of an effect

in turn, created a thicker marriage market for others. The indirect effect of a thicker marriage market for career women led even more women to opt for careers and delay marriage. An increase in the age at first marriage may also have led to higher-quality matches, if preferences are not fully formed at younger ages.

The key empirical predictions of the framework are that the introduction of the pill should have been associated with an increase in professional careers for women, the age at first marriage, and the age at first birth. Positive assortative mating on earnings capacity and compatibility among marriage partners should also have increased. The new equilibrium, however, is not completely "win-win." Women with poorer labor market prospects may suffer a decrease in their rankings as marriage partners and be the losers in the era of the pill. The framework does not produce unambiguous predictions with respect to divorce. Better matches should result from the pill, but increased career prospects for women outside marriage, decreased division of labor in the home, and potentially fewer children could all increase divorce.

Model/Framework (Cont.)

- Costa and Kahn (2003)
  - Empirical Framework right after intro
    - Pay lip service to theory papers
    - State hypotheses up front
    - Explain why think that various variables matter for production function
    - Explain where variation coming from
  - Econometric Framework after describe data
    - What regressions are you running?
    - Get to the point

Our empirical strategy uses four time-varying independent competing risk hazard models to estimate days from entry into the company (muster-in) until 1) the first case of desertion, arrest, or AWOL, 2) desertion, 3) arrests preceding desertion, and 4) AWOLs preceding desertion. We use a competing risk framework because morale varies over time, because men can become more committed soldiers, and because of censoring—some men may have died, been discharged, changed company, become prisoners of war, or be missing in action before they could desert. We treat these men as censored in our estimation strategy. When we examine time until first arrest or AWOL, we also treat men who deserted as censored (see Figure I). Note that we are assuming that the risk of desertion, arrest, or AWOL is independent of the outcomes such as death that we censor on. Hazard models provide a framework to estimate the micro and macro determinants of
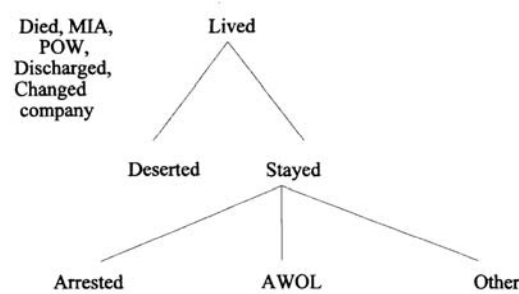
Died, MIA,
POW,
Discharged,
Changed
company

Lived

Deserted   Stayed

Arrested   AWOL   Other

FIGURE I
Schematic of Events Studied

cowardice and heroism. Our estimated hazard, $\lambda_i(t)$, for one of our four models (i), is

$$(3) \qquad \lambda_i(t) = \exp(x_I'\beta_I + x_C'\beta_C + x_D'\beta_D + x_M'\beta_M)\lambda_{i0}(t),$$

where I indexes the individual variables, C indexes the community variables, D indexes the ideology variables, M indexes the morale variables, and $\lambda_{i0}(t)$ is the baseline hazard which we assume to be Weibull. The survival function thus takes the form, $\exp((-\lambda_{ij}t_j)^p)$ for subject j, where p is the duration dependence parameter and can be interpreted as representing whether men who were in the war longer became more or less committed soldiers.[15] We present results both with and without the morale variable. The hazard ratios that we report indicate whether a one-unit change in an independent variable gives an increase/ decrease in the odds of an event. Thus, a hazard ratio of 1.3 on our Irish-born dummy variable indicates that the Irish were 1.3 times as likely as the native-born to desert. We account for unobserved company-level correlation by using variance correction models [Lee, Wei, and Amato 1992; Cai, Wei, and Wilcox 2000]. Clustering on companies provides us with an upper bound on the standard error of company characteristics.

15. Because some men may be so loyal that they would never desert, we also estimated models that account for individual heterogeneity. These yielded virtually identical results. We also tested whether censoring men who served beyond three years affected the results. We found that the magnitude of the coefficients and of the duration dependence parameter was similar, but that the standard error of some of our coefficients (e.g., occupational fragmentation, percent of the

Putting Results in Context

- This should be a whole new section
- Costa and Kahn (2003): Implications Section
  - Relate findings to literature
  - Give predicted values and show which predictors most important
  - Public policy issues: why army doesn't have 100% homogeneity today
- Not all papers will have this type of section

Some alternatives sections after the results section

- Goldin and Katz (2002): Alternative Explanations

- Katz and Murphy (1992): Understanding Changes in Education and Experience Differentials

- Goldin and Margo (1992): Explaining the Great Compression

Literature Review

- This should not be a section
- Place yourself in the literature in the introduction
- Refer to the literature in the model/framework section, the context section, and the conclusion

What goes in a footnote?

- Useful but use wisely -- you're not lawyers
- Use for details that need to be in but that you don't want cluttering up the text
- e.g. "See X for more detailed treatments on the effects of international trade flows on labor demand."
- e.g. "Katz and Murphy (1990) show that this demand measure is appropriate even in the presence of within-sector, factor neutral technical change."