# ECON4004 Econometrics 2

# Lab 2

**Question 1** (based on Stock and Watson, Exercise E8.1)

Lead is toxic, particularly for young children, and for this reason, government regulations severely restrict the amount of lead in our environment. But this was not always the case. In the early part of the 20th century, the underground water pipes in many U.S. cities contained lead, and lead from these pipes leached into drinking water. In this exercise, you will investigate the effect of these lead water pipes on infant mortality. Using the data file **lead_mortality.dta**, which contains data on infant mortality, type of water pipes (lead or nonlead, denoted by *Lead*), water acidity (*pH*), and several demographic variables for 172 U.S. cities in 1900.

**a**. Compute the average infant mortality rate (*Inf*) for cities with lead pipes and for cities with nonlead pipes. Is there a statistically significant difference in the averages?

**b**. The amount of lead leached from lead pipes depends on the chemistry of the water running through the pipes. The more acidic the water is (that is, the lower its pH), the more lead is leached. Run a regression of *Inf* on *Lead*, *pH*, and the interaction term *Lead* * *pH*.

**i**. The regression includes four coefficients (the intercept and the three coefficients multiplying the regressors). Explain what each coefficient measures.

**ii**. Plot the estimated regression function relating *Inf* to *pH* for *Lead* = 0 and for *Lead* = 1. Describe the differences in the regression functions and relate these differences to the coefficients you discussed in (i).

**iii**. Does *Lead* have a statistically significant effect on infant mortality? Explain.

**iv**. Does the effect of *Lead* on infant mortality depend on *pH*? Is this dependence statistically significant?

**v**. What is the median value of *pH* in the sample? At this *pH* level, what is the estimated effect of *Lead* on infant mortality? What is the standard deviation of *pH*?

Suppose the *pH* level is one standard deviation lower than the median level of *pH* in the sample: What is the estimated effect of *Lead* on infant mortality? What if *pH* is one standard deviation higher than the median value?

**Question 2** (based on Stock and Watson, Exercise E13.1)

A prospective employer receives two resumes: a resume from a white job applicant and a similar resume from an African American applicant. Is the employer more likely to call back the white applicant to arrange an interview? Marianne Bertrand and Sendhil Mullainathan carried out a randomized controlled experiment to answer this question. Because race is not typically included on a resume, they differentiated resumes on the basis of “white-sounding names” such as Emily Walsh or Gregory Baker) and “African American–sounding names”

(such as Lakisha Washington or Jamal Jones). A large collection of fictitious resumes was created, and the presupposed “race” (based on the “sound” of the name) was randomly assigned to each resume. These resumes were sent to prospective employers to see which resumes generated a phone call (a callback) from the prospective employer. Data from the experiment can be found in the files **names.dta**.

**a**. Define the callback rate (denoted by the variable *call_back*) as the fraction of resumes that generate a phone call from the prospective employer. What was the callback rate for whites? For African Americans? Construct a 95% confidence interval for the difference in the callback rates. Is the difference statistically significant? Is it large in a real-world sense?

**b**. Is the African American/white callback rate differential different for men than for women?

**c**. What is the difference in callback rates for high-quality versus low quality resumes? What is the high-quality/low-quality difference for white applicants? For African American applicants? Is there a significant difference in this high-quality/low-quality difference for whites versus African Americans?

**d**. The authors of the study claim that race was assigned randomly to the resumes. Is there any evidence of nonrandom assignment?