Problem Set 6: Due Wednesday, March 1.

For this problem set, choose **one** of the three problems.

Problem 1

In this problem we will sketch out how to compute elevation changes by Census blocks. Remember, these elevation changes were a key input into the housing supply elasticity constructed in Saiz (2010).

On my webpage, I have posted a python script and a do file that allows you to compute the elevation within each Census block for five counties in California. Use these files, in conjunction with USGS topographical data and data from the Census website (https://www.census.gov/geo/maps-data/data/tiger-line.html), to perform a similar analysis for a different, similarly-sized part of the United States. In particular, compute a table similar to the one that is produced in the do file (and presented at the top of the next page).

Some hints:

- The following link has a link to elevation data for the continental US:

  https://catalog.data.gov/dataset?tags=contours&res_format=TAR . Just a warning: you're downloading a somewhat large data file, here.

- From the Census website, hit the button for "2010," then click on the "Download" menu, and then click on "Web Interface." Choose "Block" from the "Select a layer type menu."

- To combine the Census and USGS data you will need to run the python code on my webpage. I would suggest using the version of python on winstat (as the necessary packages—e.g., arcpy, shutil—are already loaded).

- From here, one way to go would be compute the slopes in stata.

Problem 2

A common approach to distinguish labor supply from labor demand shocks—applied in Saiz (2010) and elsewhere (including Blanchard and Katz, 1992, and Diamond 2016, "The Determinants and Welfare Implications of US Workers' Diverging Location Choices by Skill: 1980-2000")—is to apply a Bartik instrument (so named after Bartik's 1991 book). In this problem, we will explore how this variable is constructed and the extent to which it satisfies the relevance and exogeneity conditions necessary of a good instrument.

The idea behind the Bartik instrument is to measure the change in a region's labor demand that is induced by changes in the national demand for different industries' products. The series can be constructed as follows:

$$\Delta B_{i,t-k \text{ to } t} = \sum_{I \in \text{Industries}} \underbrace{\frac{\text{emp}^I_{i,t-k}}{\sum_{I' \in \text{Industries}} \text{emp}^{I'}_{i,t-k}}}_{\text{Term (1)}} \cdot \underbrace{\log\left(\frac{\sum_{j \in \text{Regions} \neq i} \text{emp}^I_{j,t}}{\sum_{j \in \text{Regions} \neq i} \text{emp}^I_{j,t-k}}\right)}_{\text{Term (2)}} \quad ,$$

where emp^I_{it} is the employment in industry I in region i at time t. Term (1) gives the share of employment in region i that is employed in industry I in some base period. For example, for the auto industry this term will be relatively high in Detroit and relatively low in Madison. The second term gives the change in employment, in industry I, for all *other* regions. It would measure, continuing our example, how employment in the auto industry has changed in the U.S. outside of Detroit.

There are two potential data sources that you can use to compute these Bartik instruments. One is the Census County Business Patterns database, which contains info on employment in each county (or MSA) for each year between 1986 and today (the data actually go back further, but are less easily accessible). The second potential data source is the Current Population Survey which contains information on individuals' wages, the industries in which they work, and their region of residence.

In this problem, I would like you to do the following:

(a) Pick either Saiz (2010) or some other paper which uses a Bartik-type instrument. Construct the Bartik instrument as the authors do in the paper. For a few regions, plot a time series of this variable.

(b) Run a regression of changes in regional labor demand on the left hand side and ΔB on the right hand side. What is the explanatory power of this regression? In other words, how *relevant* is the Bartik instrument as a source of labor demand shocks?

(c) Make alternative Bartik instrument data series $\Delta B_{i,t-k \text{ to } t}$ by playing around with the industry definitions and period lengths. Try a coarser industry definition (if, for example, the paper you chose applies a 2-digit industry classification, then you can recompute ΔB using a 1-digit industry classification). Also, compute an alternative ΔB series with a different period length (k) than the one used in your paper. How much do your answers to (a) and (b) change with these alternative definitions?

(d) As Blanchard and Katz (1992) write regarding the Bartik instrument, "This series will be valid for our purposes as long as the national growth rates are not correlated

with labor supply shocks in the state [their definition of a region]" (page 25). Why, in the context of the paper you chose, may it be the case that regional supply shocks are correlated with the national growth rates? (Some potential regional supply shocks that you might want to think about are migration/immigration from other regions/countries or local changes in unemployment benefits..)

Problem 3

Consider the material we have covered up to now in our class. Related to any of this material, I would like you to come up with your own homework problem (along with a solution to this problem). Ideally, the scope of the homework problem that you come up with will be of the level of Problem 1 of Problem Set 1, or Problem 1 of Problem Set 4. The idea behind this exercise is for me to figure out what you find interesting, and to get you play around with data. There is the additional benefit that if you come up with an interesting problem, then I can use it in future problem sets with the knowledge that your future classmates will find your problem interesting as well. (Alternatively, if you think there is an important derivation—along the lines of Problem 1 of Problem Set 3—that we have glossed over in class, then it would be fine, too, to write and solve an exercise along those lines).