Online Appendix for "The Aggregate Implications of Regional Business Cycles"

Martin Beraja Erik Hurst Juan Ospina
University of Chicago

January 13, 2016

Preliminary

Abstract

This document is the Online Appendix that accompanies "The Aggregate Implications of Regional Business Cycles".

1 Descriptive Statistics For Retail Scanner Data

Online Robustness Appendix Table O1 shows descriptive statistics for the Nielsen Retail Scanner Database for each year between 2006 and 2012 and for the sample as a whole. A few things are of particular note. The sample sizes - in terms of stores covered - increased from 32,642 stores (in 2006) to 36,059 stores (in 2012). Second, notice that the number of observations (store*week*UPC code) is massive. The database includes over *90 billion* unique observations. Third, during the entire sample, there is about 1.4 million unique UPC codes within the database. On average, each year contains roughly 750,000 UPC codes. Fourth, the geographic coverage of the database is substantial in that it includes stores for about 80 percent of all counties within the United States. Moreover, the number of geographical units (zip codes, counties, MSAs, states) is very similar from year to year highlighting that the geographical coverage is consistent through time. Finally, the dataset includes between $188 billion and $240 billion of transactions within each year. For the time periods we study, this represents roughly 30 percent of total U.S. expenditures on food and beverages (purchased for off-premise consumption) and roughly 2 percent of total household consumption.[1]

2 Creating the Scanner Data Price Index

In this sub-section, we discuss our procedure for computing the Retail Scanner Price Indices. Formally, the first step is to produce a category-level price index which can be expressed as follows:

$$P^L_{j,t,y,k} = P^L_{j,t-1,y,k} \times \frac{\sum_{i \in j} p_{i,t,k} \bar{q}_{i,t-1,k}}{\sum_{i \in j} p_{i,t-1,k} \bar{q}_{i,t-1,k}}$$

where $P^L_{j,t,y,k}$ is price index for category j, in year t, with base year y, in geography k. For our analysis, geographies will either be U.S. states or the country as a whole. $p_{i,t,k}$ is the price at time t of the specific good i in geography k and $\bar{q}_{i,t-1,k}$ is the average monthly quantity sold of good i in the prior year in location k. By fixing quantities at their prior year's level, we are holding fixed household's consumption patterns as prices change. We update the basket of goods each year, and chain the resulting indices to produce one chained index for each category in each geography, denoted by $P^L_{j,t,k}$. In this way, the index for months in 2007 uses the quantity weights defined using 2006 quantities and the index for months in 2008 uses the quantity weights defined using 2007 quantities. This implies that the price changes we document below with changing local economic conditions is not the result of changing household consumption patterns. Fixing the basket also minimizes the well documented chain drift problems of using scanner data

[1] To make these calculations, we compare the total transcation value in the scanner data to BEA reports of total spending on food and beverages (purchased for off-premise consumption) and total household consumption.

to compute price indices (Dielwert et al. (2011)). Notice, this procedure is very similar to the way the BLS builds category-level first stage for their price indices.

When computing our monthly price indices, one issue we confront is how to deal with missing values from period to period. For example, a product that shows up in month m may not have a transacted price in month $m+1$ making it impossible to compute the price change for that good between the two months. Missing values may be due to new products entering the market, old products withdrawing from the market, and seasonality in sales. Our results in the paper are robust to the various ways we dealt with missing values but clearly the price indices will generally differ depending on how one treats such data points. Although we could have used some ad hoc imputation methods like interpolation between observed prices or keeping a price fixed until a new observation appears, we chose to follow a more conservative approach. Looking at the above equation, we see that we can handle the missing values without imputation by restricting the goods that enter the basket to those that have positive sales over at least one month in the previous year and over the 12 months of the current year. This is what we do when creating our indices. For example, when computing the category prices in 2008 we use the reference basket for 2007. In doing so, we only take the goods that have $q_{i,\bar{2007},k} > 0$ and $q_{i,t,k} > 0$ for all $t \in 2008$.[2] This ensures that for a given product in the price index during year t, we will have a weight for this product based on $t-1$ data and we will have a non-missing transaction price in all months in which the price index is computed during that year.[3] The bottom row of Appendix Table A1 includes the share of all expenditures (value weighted) that were included in our price index for a given year. In the five later years of the sample, our price index includes roughly two-thirds of all prices (value weighted).

The second stage of our price indices also follows the BLS procedure in that we aggregate the category-level price indices into an aggregate index for each location k. The inputs are the category-level prices and the total expenditures of each category. Specifically, for each state we compute:

$$\frac{P_{t,k}}{P_{t-1,k}} = \prod_{j=1}^{N} \left(\frac{P^L_{j,t,y,k}}{P^L_{j,t-1,y,k}} \right)^{\frac{\bar{S}^t_{j,k} + \bar{S}^{t-1}_{j,k}}{2}}$$

where $\bar{S}^t_{j,k}$ is the share of expenditure of category j in month t in location k averaged over the year. We calculate the shares using total expenditure on all goods in each category, even though for the category-level indices some goods were not included due to missing data. For the purposes of this paper, we make our baseline specification one that fixes the weights of each category for a year in the same fashion as we did for the

[2]The database starts in 2006. As a result, our baseline specification of the 2006 price indices only includes products that have positive sales in all months of 2006.

[3]This procedure implies that we will miss products that are introduced within a given year. These products, however, will be incorporated in next year's basket as long as they have continuous sales during the subsequent calendar year.

category-level indices. However, as a robustness specification, we allowed the weights in the second step to be updated monthly. The results using the two methods were nearly identical.

3 Creating Composition Adjusted Wage Measures in the ACS and CPS

To make the composition adjusted wage measures in the 2000 U.S. Census and the 2001-2012 American Community Survey (ACS), we start with the raw annual data files that we downloaded directly from the IPUMS website.[4] For each year, we restrict the sample to include male individuals between the ages of 25 and 55 (inclusive) who do not live in group quarters. We further restrict the sample to individuals who (1) are currently employed, (2) report working usually more than 30 hours per week (inclusive), (3) report working at least 48 weeks during the prior year, and (4) report earning at least 5000 dollars during the prior year. These latter restrictions select workers with a strong attachment to the labor force. For each individual, we create a measure of hourly wages. We do this by dividing annual labor income earned during the prior twelve month period by reported hours worked during that same time period. Our labor income also includes business income.[5] Hours worked are computed by multiplier weeks worked during the prior twelve month period times usual weekly hours worked. With the data, we compute wage measures for each year between 2000 and 2012. We wish to stress that within the ACS, the prior year refers to the prior 12 months before the survey takes place (not the prior calendar year). Individuals interviewed in January of year t report earnings and weeks worked between January and December of year $t-1$. Individuals in June of year t report earnings between June of year $t-1$ and May of year t. Given that the ACS samples individuals in every month, the wage measures we create for year t can be thought of as representing average wages between the middle of year $t-1$ through middle of year t. This differs slightly from the timing in the Current Population Survey (CPS) which we discuss next.

To create measures of composition adjusted wages, we regress $\ln(wage_{it})$ on three dummies for usual weekly hours worked (hours worked 30-39, hours worked 50-59, and hours worked 60+), five age dummies (age 25-29, age 30-34, age 35-39, age 45-49, and age 50-55), four education dummies (less than high school, exactly high school, exactly a bachelors degree, and more than a bachelors degree), a dummy if the individual's race is black, and dummies for the individual's citizenship. This regression is run separately for each year. We weight each regression using the individual survey weights provided by the Census and ACS. After running the regression, we compute the residuals for each individual. To rescale the residuals, we add back in the regression constant for all

[4]The ACS is just the annual survey which replaces the Census long form in off Census years. The national representative survey started in 2001. As a result, the Census and ACS questions are identical.

[5]Our results were robust to excluding business income from our measure of labor earnings.

individuals. Finally, we convert the composition adjusted log wage residuals back to levels. To make the annual state level composition adjusted wage index, we just take the weighted average of wage residuals across individuals in each state separately for each year.

To examine longer aggregate trends in composition adjusted wages, we use data from the March Current Population Survey. We download the data directly from the IPUMS website. As with the ACS data, we restrict the sample to men between the ages of 25 and 55 who do not live in group quarters and who have a strong attachment to the labor force (currently employed, worked 48 weeks during the prior year, and currently report working at least 30 hours per week during a usual week). We also restrict the data to those individuals with positive sample weights. Our procedure for making composition adjusted wages are identical to above aside from the following three changes: (1) we do not control for citizenship given that citizenship is not consistently measured over our sample period, (2) our sample period is 1977 through 2012, and (3) we pool years together when running regressions to increase power. In particular, we run our regression on three separate time periods: 1977-1995, 1996-2000, and 2000-2012. We break the sample into three periods because the CPS earnings questions changed in 1995 and the CPS expanded its sample in 2000.

4 BLS Metro Area Price Indices

The U.S. Bureau of Labor Statistics produces 27 metro areas price indices at various degrees of time aggregation.[6] For each metro area, the BLS publishes multiple price indices (food, nondurables, etc.). Although there are only 27 MSAs, we can still explore the relationship between unemployment growth between 2007 and 2010 and the cumulative inflation rate between 2007 and 2010 using these data. We then can compare the relationship between local unemployment growth and local scanner price inflation that we document in Section 3 with the relationship between local unemployment growth and local inflation computed using the BLS data. Given the price indices are only provided semi-annually for many MSAs, we compare the change in the unemployment rate and prices from the first half of 2007 to the latter half of 2010. When data is provided monthly or bi-monthly, we simply take the geometric average over the first and last half of the year to make the data semi-annual. As with the results in the main paper, we use data from the BLS's local area unemployment statistics to measure the percentage point change in the local unemployment rate.

Using this data, we regress the 3-year inflation rate at the MSA level on the 3 year change in the unemployment rate. We run this regression for various inflation measures

[6]The 27 MSAs are (in order of reporting frequency): Chicago, Los Angeles, New York, Atlanta, Boston, Cleveland, Dallas-Fort Worth, Detroit, Houston, Miami, Philadelphia, San Fransicso, Seattle, Washington, Anchorage, Cincinnati, Denver, Honolulu, Kansas City, Milwaukee, Minneapolis, Phoenix, Pittsburg, Portalnd, St. Louis, San Diego, and Tampa. The first three MSAs have price indices that are reported at monthly frequencies. The next 11 MSAs have price indices that are reported at bi-monthly frequencies. The last 13 price indices are reported semi-annually.

(food, services, all goods less housing, etc.). The results of these simple regressions using the BLS data are strikingly similar to our results using the scanner data. For example, the results using the scanner index find that a 1 percentage point increase in the local unemployment rate is associated with a 0.38 decline in the local inflation rate (for the specification where goods are defined as UPC-store pairs). Within the BLS data, we find that a 1 percentage point increase in the local unemployment rate is associated with a 0.34 percentage point decline in the local food inflation rate (standard error = 0. 22).

Also, as predicted by our simple model in the paper, the relationship between the inflation rate for all goods and the local unemployment rate change should be higher than the relationship between food inflation and the local unemployment rate change if food is relatively more tradable than the local consumption good. We cannot test this within our scanner data. However, the BLS data allows us to test this prediction directly. Within the BLS data, the relationship between the inflation rate for all goods with the change in the unemployment rate is in fact higher at -0.47 (standard error = 0.15). The fact that the coefficient is larger in magnitude is consistent with our belief that the variation in packaged food data across regions should be a lower bound for the variation in the average consumption good given that the packaged goods in our dataset are relatively more tradable.

The patterns that we document for the 2007-2010 period with the BLS data are very consistent with the results in Fitzgerald and Nicolini (2014) that use the BLS MSA level price indices to show these relationships over a much longer time period. Fitzgerald and Nicolini (2014) find that over the period of 1976-2010, a 1 percentage point increase in the local unemployment rate is associated with a 0.3 percentage point decline in the local annual inflation rate. In summary, a limitation of our price data is that it only covers goods sold in grocery, pharmacy and mass-merchandising stores. The fact that the patterns we uncover are nearly identical for similar goods using the BLS metro area price indices is very reassuring. Moreover, as predicted, using broader measures of goods in the BLS data seems to only strengthen the cross-region variation. Our data is an advance over the BLS metro level data in that can be calculated for every state and, if one desires, a much larger set of metro areas.

5 Accounting for Measurement Error in Wages and Prices

Formally, in Section 6 of the main paper, we estimate the following specification using our regional data to obtain estimates of λ and ϕ:

$$\pi^w_{kt} = b_0 + b_1 \pi_{kt} + b_2(n_{kt} - n_{k,t-1}) + b_3 \pi^w_{k,t-1} + \Psi D_t + \Gamma X_k + e_{kt}$$

where $b_1 = \lambda$, $b_2 = \lambda/\phi$, $b_3 = (1-\lambda)$, and $b_t = \lambda(u^\epsilon_t - (1-\rho_\epsilon)\epsilon_{t-1})$. However, given our data construction procedures, the regional inflation rate (π_{kt}), regional nominal wage growth (π^w_{kt}), and regional lagged nominal wage growth (π^w_{kt-1}) are all measured with error. If the measurement error were classical, our estimates of b_1 and b_3 would be attenuated. Moreover, given that we measure nominal wages in each period and then compute the changes over time to get the growth rates, any measurement error in our nominal wage measure in a given year will induce a negative correlation between π^w_{kt} and π^w_{kt-1}. This type of measurement error would cause b_3 to be biased downward.

Given that we want to recover structural parameters, we take these measurement errors seriously. To account for the measurement error in π_{kt}, π^w_{kt}, and π^w_{kt-1}, we exploit the large sample sizes underlying our the construction of local price and wage measures. We begin by addressing the measurement error in prices. Using the underlying data in from Nielsen, we split the data in half by product categories. As discussed, the Nielsen data includes roughly 1,000 product categories. We split the underlying data into two groups of categories: sample 1 includes all the odd category numbers (1, 3, 5, etc.) while sample 2 includes all the even category numbers (2, 4, 6, ect.). We then construct state level prices indices for each month using the data from sample one (P^{samp1}_{kt}) and then separately using the data from sample two (P^{samp2}_{kt}). When running our key equations, we instrument the inflation rate using the price indices computed with data from sample 1 (π^{samp1}_{kt}) with the inflation rate using the price indices computed with the data from sample 2 (π^{samp2}_{kt}). Not surprisingly, the annual inflation rates obtained from each of the two separate data samples are highly correlated. Specifically, running the first stage equation on our annual data from 2007-2011:

$$\pi^{samp1}_{kt} = \psi_0 + \psi_1 \pi^{samp2}_{kt} + \psi_2 D_t + \psi_{error}$$

yields an estimate of ψ_1 = 0.268 (standard error = 0.06) with an adjusted R-squared of 0.84. The F-stat from including π^{samp2}_{kt} is 19.6. Using the above relationship, we can make a predicted inflation measure:

$$\hat{\pi}_{kt} = \hat{\psi}_0 + \hat{\psi}_1 \pi^{samp2}_{kt} + \hat{\psi}_2 D_t$$

It is this predicted inflation measure that we include in our estimating equation. We adjust standard errors accordingly to account for the predicted regressor.

We perform a similar methodology to adjust for the measurement error in wages. Specifically, we use a random number generator to split the underlying micro data from the ACS into two equal sized samples. Individuals within a year with a random number less than or equal to 0.5 go into sample 1 and individuals within a year with a random number greater than 0.5 go into sample 2. As a result, for each year, we will have three samples of data within each year: sample 1 individuals, sample 2 individuals, and all individuals. Within each sample, we can make measures of adjusted nominal wages (where the adjustment for observables as is discussed in the paper). With the adjusted nominal wages for each year and each sample, we can compute nominal wage growth for each year-sample pair. To account for the measurement error in nominal wage growth, we run the following two "first stage" equations using our annual data from 2007-2010:

$$\pi^w_{kt} = \psi^1_0 + \psi^1_1 \pi^{w,samp1}_{kt} + \psi^1_2 D_t + \psi^1_{error}$$

$$\pi^w_{kt-1} = \psi^2_0 + \psi^2_1 \pi^{w,samp2}_{kt-1} + \psi^2_2 D_t + \psi^2_{error}$$

Specifically, we use sample 1 data to predict contemporaneous wage growth and use sample 2 data to predict lagged wage growth. This ensures that any measurement error in our current wage growth measure will not be correlated with the measurement error in our lagged wage growth measure. Like with the price data, the wage data are highly correlated across the samples. For the contemporaneous wage growth equation, our estimate of ψ^1_1 is 0.80 (with a standard error of 0.03) and an adjusted R-squared of 0.89. The F-stat from including $\pi^{w,samp1}_{kt}$ is 783. For the lagged wage growth equation, our estimate of ψ^2_1 is 0.30 (with a standard error of 0.04) and and adjusted R-squared of 0.66. The F-stat from including $\pi^{w,samp2}_{kt-1}$ is 75.6. Using these regressions, we can make measurement error adjusted current and lagged wage growth measures:

$$\hat{\pi}^w_{kt} = \hat{\psi}^1_0 + \hat{\psi}^1_1 \pi^{w,samp1}_{kt} + \hat{\psi}^1_2 D_t$$

$$\hat{\pi}^w_{kt-1} = \hat{\psi}^2_0 + \hat{\psi}^2_1 \pi^{w,samp2}_{kt-1} + \hat{\psi}^2_2 D_t$$

It is these predicted measures that we use in our estimation equations. Specifically, our key estimating equation can be expressed as:

$$\hat{\pi}^w_{kt} = b_0 + b_1 \hat{\pi}_{kt} + b_2(n_{kt} - n_{k,t-1}) + b_3 \hat{\pi}^w_{k,t-1} + \Psi D_t + \Gamma X_k + e_{kt}$$

In the main text of the table, we suppress the hats on the variables. We adjust the standard errors by bootstrapping to account for the fact that the measures of price and wage growth are estimated.

6 Variation in Government Policy Changes Across States During Great Recession

In the main text, we argued that we could identify unbiased estimates of λ and ϕ using regional variation via OLS if $v^{\epsilon}_{kt} = 0$. This assumption would be violated if government policy changed differentially across states in a way that discouraged labor supply. In this section of the appendix we show that the difference in many such policy changes across states were small (particularly relative to the aggregate changes) and that these policy changes - to the extent that they did occur - occurred after 2009. Specifically, we focus our attention on four such government policies: state income tax rates, federal food assistance programs, federal programs to help underwater homeowners renegotiate their mortgage contract, and the extension of unemployment benefits. Our analysis shows that for these major policies, our assumption that these policies varied little across regions and the extent to which they did vary was uncorrelated with local measures of economic activity is not at odds with the data.

Using data on statutory tax rates by state from the tax foundation and micro data on individual incomes by state from the American Community Survey, we compute the average marginal tax rate for each state in 2007 and 2010. State income tax changes were not that common during this time period. Roughly 90 percent of the states, population weighted, had essentially no change in their average marginal tax rate during this time period. As seen from Online Appendix Figure O1, the extent to which the average marginal tax rate changed between 2007 and 2010 was uncorrelated with state employment or price growth between 2007 and 2010.

Likewise, Online Appendix Figure O2 shows no correlation in the growth in benefits from the federal Supplemental Nutrition Assistance Program (SNAP) and employment growth across states during the 2007-2010 period. SNAP is the successor to the federal Food Stamps program. This program was expanded dramatically during the Great Recession. Given that the SNAP program is means tested, an expansion of the program can discourage work effort. This point is made by Mulligan (2012). Using summary data from the US Department of Agriculture, we measure the dollar per SNAP recipient for each state between 2007 and 2010. The average recipient received an increase about 33 percent during the 2007 to 2010 period. Yet, there was very little regional variation in the increase (a standard deviation of only 4 percent across the states, population weighted). As seen from Online Appendix Figure O2, the variation that did occur across states was uncorrelated with state unemployment growth. So, while the increase in SNAP benefits may have discouraged labor supply at the aggregate level, there is very little variation across U.S. states.[7] This is not to say that SNAP total dollars paid did not differ across states. It is saying that the dollar per recipient did not vary across states. If a local shock hit that resulted in people not working, the number of people eligible for SNAP would increase. We are showing that all the variation in SNAP dollars paid across states was due to the changing number of recipients across states NOT the change in the dollar per recipient.

Another new federal program that was means tested and was argued to possibly discourage work effort during the latter part of the Great Recession was the Home Affordable Modification Program (HAMP). HAMP was designed to help homeowners who were underwater renegotiate their mortgage. The program was authorized in early 2009 but there were no significant modifications taking place until mid 2010. By the end of 2010, only about 0.5% of households had participated in the program. As seen from Online Robustness Appendix Figure O3, the extent to which households participated in the program varied slightly with the state's employment growth between 2007 and 2010. A one percent decline in employment growth was only associated with a 0.07 percentage

[7] We also explored whether the eligibility for SNAP differed across states in a way that is correlated with the change in state economic conditions between 2007 and 2010. We found no evidence suggesting such a relationship.

point increase in HAMP take up (i.e., from 0.50% to 0.57%). This effect is very small. However, given that there was some relationship between HAMP take up and underlying economic conditions within the state, we perform a robustness specification in our cross sectional estimation that excludes 2009 and 2010 data and therefore only focuses on the periods before HAMP went into effect. As discussed in the main text, these estimates were very similar to our baseline estimates. Additionally, we excluded states with the highest amount of loan modifications (CA, FL, NV and AZ) as a robustness specification. Our results were unchanged when these states were excluded.

Finally, we explored the extent to which unemployment benefits were differentially extended at the state level. During the 2007-2011 period, unemployment benefits increased from about 26 weeks per recipient to upwards of 99 weeks per recipient in some states. Some researchers have argued that this large increase in the duration of unemployment benefits can explain only a small portion of the aggregate decline in employment (Rothstein (2012)) while others have argued that it can explain a more substantive portion of the aggregate decline in employment (Hagedorn et al. (2013)). By law in 2010, weeks of unemployment benefits were tied to the state's unemployment rate. This implies that there will be a correlation between the total amount of unemployment benefit extension within the state and the change in the state employment rate between 2007 and 2010. Online Appendix Figure O4 shows this correlation. As of 2010, 70 percent of U.S. states had a duration of unemployment benefits that exceeded 86 weeks. These states represent roughly 90 percent of the U.S. population. However many smaller states, mostly in the Plains region of the U.S., had small employment declines and only an extension of unemployment benefits from 60-85 weeks.[8]

A simple regression line through Online Robustness Figure O4 shows that a 1 percent decline in employment growth was associated with an additional 1.5 weeks of unemployment benefits. Again, this is a tiny change. However, as discussed in the main text, we can reestimate our model focusing only on data before the unemployment benefit extension occurred (i.e., prior to 2010). As we show in the main text, the results are nearly identical to our base specification. Additionally, we can include only those states that had an increase in unemployment benefits to at least 86 weeks (using all years of our data).[9] If pooling together states that had large and small unemployment benefits were biasing our estimates of λ and ϕ, our estimates would change once we excluded the low unemployment benefit during states. It is comforting that our estimates did not change at all when these states were excluded. This says that the fact that unemployment benefit durations changed differentially across states is not biasing our results in any substantive way. This is not surprising when one realizes that essentially all states (population weighted) had increases in unemployment benefit duration to at least 86 weeks.

[8]States also had some discretion as to whether they opted into the program. This explains why some states did not have the maximum weeks of unemployment benefits even when their unemployment rate was higher.

[9]The excluded states are: Arkansas, Iowa, Louisiana, Maryland, Mississippi, Montana, Nebraska, New Hampshire, North Dakota, Oklahoma, South Dakota, Utah, and Wyoming.

7 Controlling for Industry Controls in Estimates λ and ϕ

Charles et al. (2013) document that the secular decline in manufacturing depressed employment rates during the 2000s. Autor and Dorn (2013) show that declines in routine employment also depressed employment rates during the 2000s. Both of these prior papers exploit variation across either MSAs or commuting zones. As we show in Online Robustness Appendix Figure O5, there is very little variation in routine occupation shares across U.S. states as of 2007 and the extent to which there is variation, it is uncorrelated with employment growth during the 2007-2010 period. We define routine occupations to include all manufacturing and administrative occupations. Although not shown, a similar pattern exists for just manufacturing employment as of 2007. This is consistent with the results in Charles et al. that most of the decline in manufacturing employment within the U.S. during the last 15 years took place in the early 2000s. As a robustness exercise, we control for the 2007 share of routine jobs within each state and the 2007 share of manufacturing jobs in each state. As seen from the main text, adding these controls do not alter our estimates in any meaningful way. We also explored adding more detailed industry controls (at the one digit level). Again, these controls did not alter our estimates of λ and ϕ in any meaningful way.

8 IV Procedure to Estimate λ and ϕ

As a specification check on our estimation procedure, we attempted to isolate variation in labor demand across states that were orthogonal to shocks to labor supply. We build on the work of many others (including Mian and Sufi 2014) by looking at changes in housing prices. Instead of instrumenting housing prices with local housing supply elasticities to isolate a causal effect of housing price changes, we use housing prices directly as a instrument. We are not interested in getting a causal effect of housing prices on local labor markets. Instead, we are interested in isolating movements in wage growth, price growth, and employment growth that is orthogonal to changes in the taste for leisure at the local level. The key condition we need is that (1) the house price changes were not caused by changes in the taste for leisure and (2) the house price changes did not cause a change in the taste for leisure, our identification strategy will be valid. Like with our OLS estimates, we estimate the IV procedure over two time periods - the full 2007-2010 period and the shorter 2007-2009 period (prior to the potentially reactionary change in state policy variables).

Specifically, we use two instruments: contemporaneous and lagged housing price growth. We instrument for contemporaneous employment growth and price growth. Given we restrict the coefficient on lagged wage growth to be one minus the coefficient on price growth, we only need two instruments to estimate λ and ϕ. Not surprisingly, the housing variables predict both employment growth and lagged wage growth. The first stage F-stat for the instruments in predicting employment growth is 17.3.

9 Parameter Calibration

In this section we explain how we determined the values of the following parameters used in the paper (see Table 4 in the paper): ρ_z (persistence of productivity shock), ρ_ϵ (persistence of the labor supply shock), and β (intermediates share in the non-tradable sector).

9.1 Persistence Parameters in Shock Processes

To measure the persistence parameters of the productivity/mark-up shock and labor supply/leisure shock we use two equations from the model from which we can back them out. For the productivity/mark-up process, z_t, we use the aggregate labor demand equation:

$$w_t^r = -(1-(\alpha+\theta\beta))n_t + z_t$$

For the labor supply/leisure process, ϵ_t, we use the aggregate wage-setting equation:

$$\pi_t^w = \frac{\lambda}{1-\lambda}(\epsilon_t + \frac{1}{\phi}n_t - w_t^r)$$

In these equations, we plug in the values for the other parameters and the aggregate data that we used in our VAR and solve for z_t and ϵ_t. With the time series of these processes we measured their AR(1) persistence coefficient obtaining $\rho_z = 0.76$ and $\rho_\epsilon = 0.66$

9.2 Input Shares

Here we just explain how we computed the values for β, the share of intermediate inputs. The labor shares in the non-tradable and tradable sector are, respectively:

$$\alpha = \frac{wN^y}{PY^y}$$

$$\theta\beta = \frac{wN^x}{PY^y}$$

The aggregate employment share in the economy is

$$\vartheta = \frac{wN^x + wN^y}{QY^x + PY^y}$$

Divide through by PY^y

$$\vartheta = \frac{\frac{wN^x}{PY^y} + \frac{wN^y}{PY^y}}{\frac{QY^x+PY^y}{PY^y}} = \frac{\theta\beta+\alpha}{1+\frac{QY^x}{PY^y}}$$

Rearranging we get β.

$$\theta\beta + \alpha = \vartheta\left(1 + \frac{QY^x}{PY^y}\right)$$

$$\theta\beta + \alpha = \vartheta\,(1 + \beta)$$

$$\beta = \frac{\vartheta - \alpha}{\theta - \vartheta}$$

We used values for employment shares in the economy, the tradable and non-tradable sector in 2006 to get a value for β of 0.16.

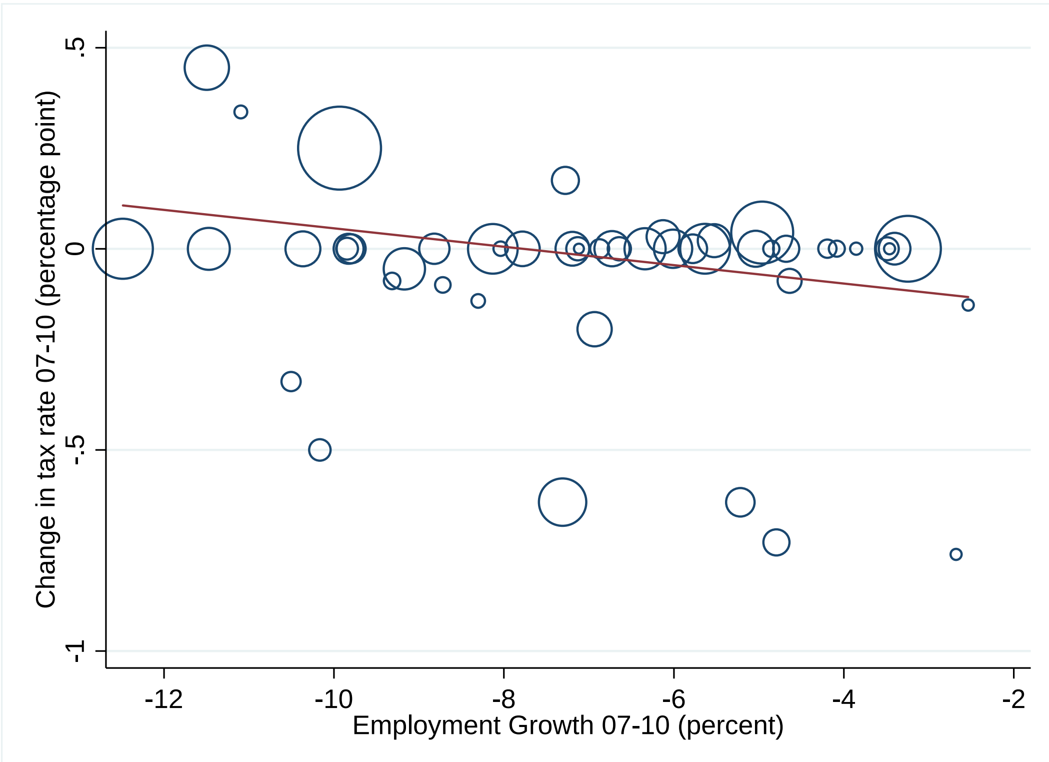
10 Figures and Tables

Table O1: Descriptive Data for the Nielsen Scanner Price Data

	Individual Years							Combined Years	
	2006	2007	2008	2009	2010	2011	2012	Total	Average
Number of Obs. (million)	12,013.1	12,812.2	13,037.5	12,968.3	13,153.4	13,646.7	13,618.8	90,250.1	13,035.7
Number of UPCs	725,224	762,469	759,989	753,984	739,768	742,074	753,318	1,425,484	748,118
Number of Chains	86	85	87	86	86	86	82	88	85
Number of Stores	32,642	33,745	34,830	35,343	35,807	35,645	36,059	39,368	34,867.3
Number of Zip Codes	10,869	11,123	11,357	11,476	11,589	11,639	11,626	11,797	11,382.7
Number of Counties	2,385	2,468	2,500	2,508	2,519	2,526	2,547	2,568	2,493.3
Number of MSAs	361	361	361	361	361	361	361	361	361
Number of States	49	49	49	49	49	49	49	49	49
Transaction Value (US billion)	187.9	207.8	219.6	223.7	227.6	235.2	239.5	1541.2	220.2
Pct. Value used in Price Index	54.3%	50.0%	66.4%	66.0%	68.3%	68.0%	67.7%	63.4%	63.0%

Note: Table shows descriptive statistics for the underlying data that we used to create our Scanner Price Index using the Nielsen Retail Scanner Data.

Figure O1: Change in State Tax Rate vs. State Employment Growth: 2007-2010

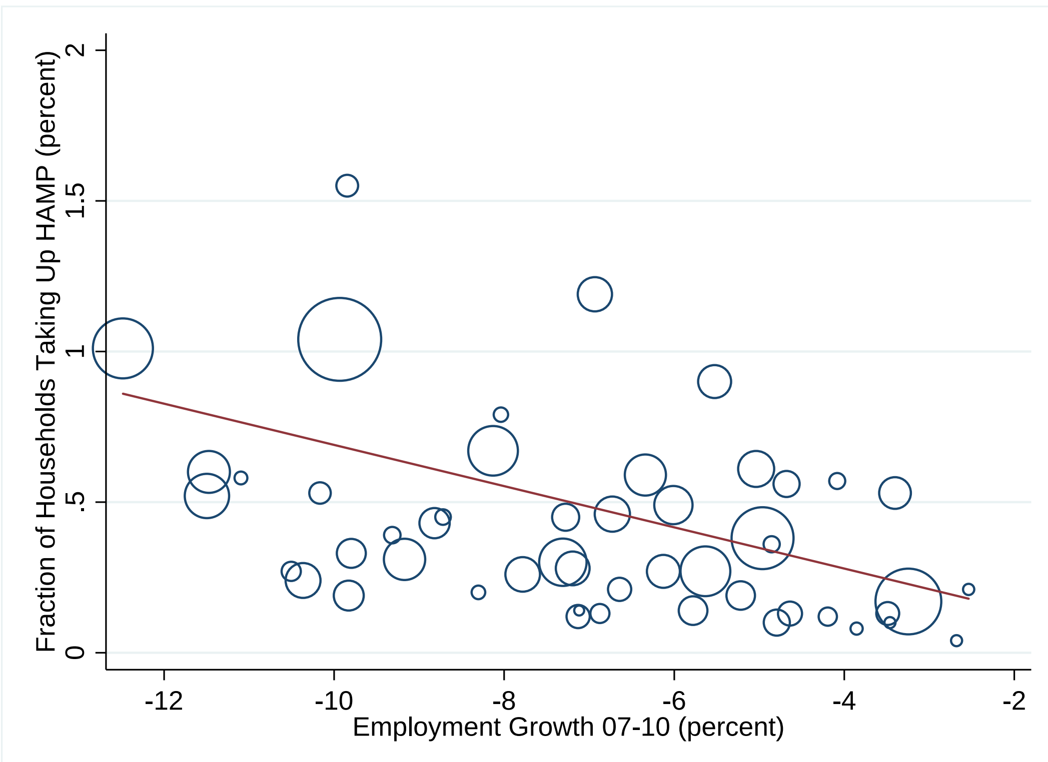


Note: Figure shows the change in the average marginal tax rate in a state between 2007 and 2010 against employment growth in the state during the same period. Employment growth comes from the BLS and is defined in the text. To compute the average marginal tax rate in the state we use data from the American Community Survey and state tax rate formulas from taxfoundatoin.org. Using the American Community Survey, we compute the fraction of state residents in 15 labor income bins as well as the mean income within each bin. We then compute the marginal tax rate in that bin. Averaging over the bins, we get the stateÕs average marginal tax rate. Our procedure does not account for any state level deductions or exemptions. Additionally, it assumes no one files jointly. It is mean to give a summary statistic for the stateÕs average marginal tax rate.

Figure O2: State SNAP Growth vs. State Employment Growth: 2007-2010

Note: Figure shows the change in SNAP payment growth per recipient at the state level between 2007 and 2010 against employment growth in the state during the same period. Employment growth comes from the BLS and is defined in the text. SNAP growth per recipient was collected from http://www.fns.usda.gov

Figure O3: State HAMP Take-Up vs. State Employment Growth 2007-2010

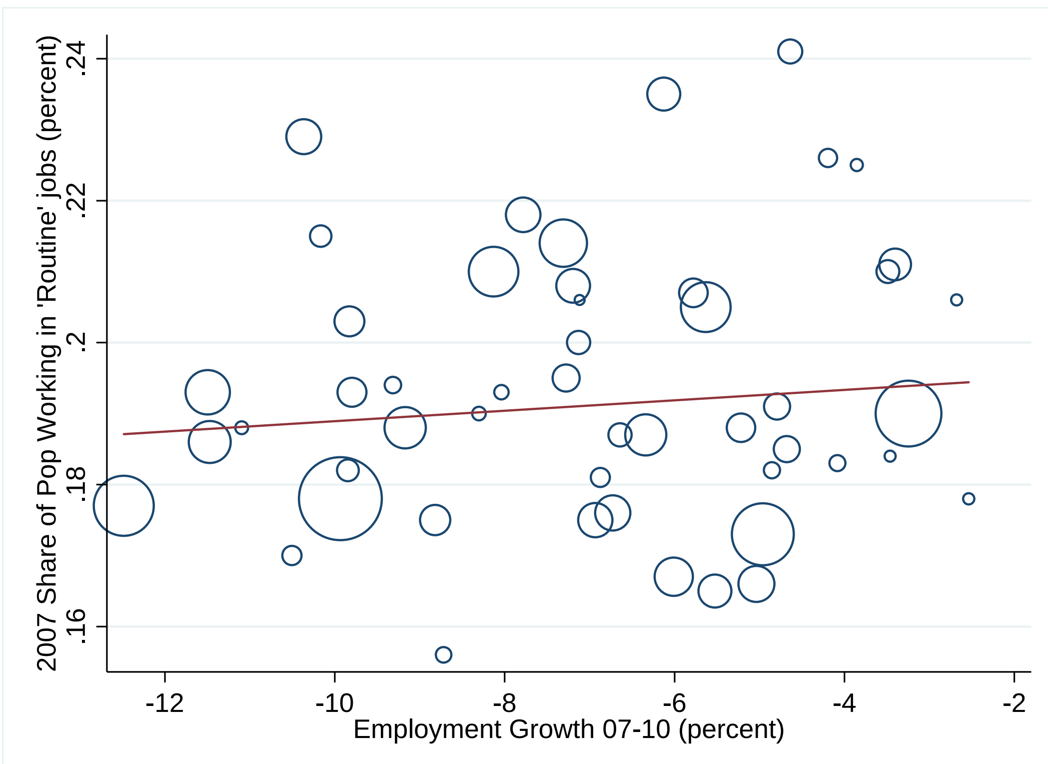


Note: Figure shows number of households participating in HAMP programs in 2010 against employment growth in the state during 2007-2010. Employment growth comes from the BLS and is defined in the text. HAMP participation comes from http://www.treasury.gov.

Figure O4: Max Unemployment Benefit Receipt in 2010 vs. State Employment Growth 2007-2010

Note: Figure shows the maximum number of unemployment benefits allowed in state in 2010 against employment growth in the state during 2007-2010. Employment growth comes from the BLS and is defined in the text.

Figure O5: Routine Share of Employment in 2007 vs. State Employment Growth 2007-2010



Note: Figure shows the routine share of employment in the state in year 2007 against employment growth in the state during 2007-2010. Routine employment is defined as anyone working in a manufacturing or administrative job. Routine employment shares are computed from the 2007 American Community Survey.