On line Appendix: A. Data appendix

Descriptive Statistics

Table A.1: Distribution of annual household income in BRFSS 2005-2009

Item	Midpoint	Number	Per cent
Less than $10,000	$7,954	63,625	4
Between $10,000 and $15,000	$12,828	66,513	4
Between $15,000 and $20,000	$17,534	93,129	6
Between $20,000 and $25,000	$23,102	115,946	7
Between $25,000 and $35,000	$30,718	155,076	10
Between $35,000 and $50,000	$42,767	207,710	13
Between $50,000 and $75,000	$67,718	240,541	15
Greater than $75,000	$125,362	445,086	28
Information is missing		179,524	11
Total		1,567,150	100

Notes: (1) The table shows statistics from the BRFSS sample used in estimations; about 300,000 observations were omitted, mostly because those respondents did not supply information on county of residence or because we do not have unemployment statistics for their counties of residence. (2) Analytical weights are used. (3) The midpoint is approximated under the assumption that household income follows a lognormal distribution. The mean of that distribution is observed from the survey itself (excluding those missing income information); the standard deviation is the average of all imputed values based on individual cutoff points (the cutoff point is discarded if it coincides with the mean).

Table A.2: Distribution of monthly household income in Gallup Daily Poll 2008-2009

Monthly income	Midpoint(annual)	Number	Per cent
$500 to $999	$9,818	38,020	6
$1,000 to $1,999	$19,576	84,459	13
$2,000 to $2,999	$31,287	77,334	12
$3,000 to $3,999	$43,020	64,263	10
$4,000 to $4,999	$55,603	53,638	8
$5,000 to $7,499	$77,135	89,131	13
$7,500 to $9,999	$110,860	37,733	6
$10,000 and over	$180,338	68,968	10
Information is missing		147,784	22
Total		661,330	100

Notes: (1) The table shows statistics from the Gallup Daily Poll's sample used in our estimations; about 20,000 observations were omitted, mostly because those respondents did not supply information on county of residence or because we do not have unemployment statistics for their counties of residence. (2) Analytical weights are used. (3) The midpoint is approximated under the assumption that household income follows a lognormal distribution. The mean of that distribution is observed from the survey itself (excluding those missing income information); the standard deviation is the average of all imputed values based on individual cutoff points (the cutoff point is discarded if it coincides with the mean).

Table A.3: Summary statistics for other variables in BRFSS 2005-2009

Variable	**Mean**	**Std. Dev.**	**Min.**	**Max.**	**N**
UR: unemp. rate in county	0.06	0.03	0.01	0.32	1567150
LFS: Unemployed	0.06	0.24	0	1	1567150
LFS: Not in labor force	0.33	0.47	0	1	1567150
Male	0.49	0.5	0	1	1567150
Age 18 to 29	0.2	0.4	0	1	1567150
Age 30 to 49	0.38	0.48	0	1	1567150
Age 50 to 64	0.24	0.43	0	1	1567150
Age 65 or above	0.17	0.37	0	1	1567150
Edu: High sch. or below	0.38	0.49	0	1	1567150
Edu: Some post-secondary	0.27	0.44	0	1	1567150
Edu: University degree	0.35	0.48	0	1	1567150
Single/never married	0.19	0.39	0	1	1567150
Married/with partner	0.64	0.48	0	1	1567150
Divorced/seprt./widowed	0.17	0.38	0	1	1567150

Continued on next page...

... table A.3 continued

Variable	Mean	Std. Dev.	Min.	Max.	N
Race: White	0.70	0.46	0	1	1567150
Race: Black	0.1	0.3	0	1	1567150
Race: Hispanic	0.14	0.34	0	1	1567150
Race: Others	0.06	0.25	0	1	1567150

Notes: (1) The table shows statistics from the BRFSS sample used in our estimations; about 300,000 observations were omitted, mostly because those respondents did not supply information on county of residence or because we do not have unemployment statistics for their counties of residence. (2) Analytical weights are used.

Table A.4: Summary statistics for other variables in Gallup Daily Poll 2008-2009

Variable	Mean	Std. Dev.	Min.	Max.	N
Proxy for u-index	0.11	0.31	0	1	649687
UR: unemp. rate in county	0.08	0.03	0.01	0.32	661330
Male	0.48	0.5	0	1	661330
Age 18 to 29	0.17	0.38	0	1	661330
Age 30 to 49	0.35	0.48	0	1	661330
Age 50 to 64	0.26	0.44	0	1	661330
Age 65 or above	0.19	0.39	0	1	661330
Edu: High sch. or below	0.47	0.5	0	1	661330
Edu: Some post-secondary	0.22	0.42	0	1	661330
Edu: University degree	0.3	0.46	0	1	661330
Single/never married	0.2	0.4	0	1	661330
Married/with partner	0.6	0.49	0	1	661330
Divorced/seprt./widowed	0.2	0.4	0	1	661330
Race: White	0.75	0.43	0	1	661330
Race: Black	0.09	0.29	0	1	661330
Race: Hispanic	0.11	0.31	0	1	661330
Race: Others	0.05	0.21	0	1	661330

Notes: (1) The table shows statistics from the Gallup Daily Poll's sample used in our estimations; about 20,000 observations were omitted, mostly because those respondents did not supply information on county of residence or because we do not have unemployment statistics for their counties of residence. (2) Analytical weights are used.

Data Sources

Data sources for unemployment insurance. The ratio of average weekly benefit to average weekly total wage is from the CLAIMS DATA REPORT (Taxable and Reimbursable) in the Annual Program and Financial Data (ET Financial Handbook 394) available from the Employment and Training Administration at the Department of Labor.[12] The data between 2005 and 2008 are from the annual reports. We also backed out the average weekly wage by dividing the average weekly benefit amount by the ratio. The annual report for 2009 has yet to be released as of June 6th, 2010. The average replacement ratios are very stable within a state over the period between 2005 and 2008. We assume the stability extends to the 2009, and use the 2008 in place of the 2009 data. The maximum benefit amount by states between 2005 and 2009 is from the annual "Comparison of State Unemployment Insurance Laws" at the same source (URL: http://www.ows.doleta.gov/unemploy/comparison2009.asp). The maximum amount in a number of states is a range instead of a single value; we use the upper limit.

B. Constructing the instrumental variable for local unemployment rates

For the instrument approach, we compute a series of annual likely employment loss for each county as a faction of that county's total employment. The likely loss is calculated based on the county's composition of employees by industries, together with the state-wide employment loss by industries. More specifically, the likely loss between quarter $t-4$ and quarter t for

[12]URL: http://www.ows.doleta.gov/unemploy/finance.asp. The so-named 'Reimbursable' include public employees and non-for-profit employees

county j in state s, as a fraction of employment stock in the county, is given by

$$LikelyEmpLossRate_{j,s,t} = \frac{\sum_{x=1}^{X}(lnN_{x,s,t-4} - lnN_{x,s,t}) * N_{x,j,s,t}}{\sum_{x=1}^{X} N_{x,j,s,t}}$$

On the right-hand side of the equation, the denominator is the total employment in the county. The numerator is the likely employment loss during the four quarters from t-4 to t; it is the sum of likely employment losses from all sectors in that county, with sectors denoted as x in subscripts. For a specific sector in a specific county, the likely employment loss is the product of two factors: one is the proportional employment loss of that sector, approximated by difference in logs, in the entire the state that the county belongs to; the other is the number of workers in the county who are working in that particular sector. We work on the "Supersectors" as defined in the Quarterly Census of Employment and Wage (QCEW) program. These sectors are construction, education and health services, financial activities, information, leisure and hospitality, manufacturing, natural resources and mining, other services, professional and business services, trade, transportation, and utilities, and the unclassified. All data are from the QCEW statistics available on line from the BLS.

A positive value of this $LikelyEmpLossRate_{j,s,t}$ indicates a negative labor market shock to the county involved; it is calculated from industrial composition of the county and state-wide events; we are reasonably confident that it captures exogenous shocks to the county. We will use the series of likely losses to predict the current levels of unemployment rates using a fixed-effect panel regression. In order to remove anything that may be re-

lated to local culture and others, we remove county-specific intercepts from the predicted values. The fitted value, net of county-specific intercepts, are then used to replace the actual unemployment rates in Table 1.

The first-stage estimates are shown in Table A.5. The fixed-effect panel has counties on the spatial dimension and quarters on the time dimension. The dependent variable is the level of unemployment rates. The explanatory variable include county-specific intercepts, the likely employment loss rate in the past one year and the lagged values of these likely losses in an interval of four quarters; four such lags are included. The sample period is between 2005 and 2009; with four-year lags, however, the explanatory variable goes back to 2001. Finally, we use all the available counties for the panel regression, even if some of the counties are not included in the second stage of the analysis.

Table A.5: The first stage of the instrumental variable approach: predicting local unemployment rates using current and lagged likely employment losses

	unempRate
Variables	(1)
Likely employment loss rate	46.48
	(0.26)***
L4.likelyEmpLossRate	20.39
	(0.47)***
L8.likelyEmpLossRate	12.97
	(0.45)***
L12.likelyEmpLossRate	12.96
	(0.41)***
L16.likelyEmpLossRate	-4.72
	(0.42)***
Obs.	62394
R^2	0.63
F statistic	20401.14

Notes: (1) the variables shown on the top row are dependent variables. (2) The numbers in the parentheses are standard errors. (3) *, **, and *** indicate statistical significance at 10%, 5% and 1% levels. (4) Sample period is between the first quarter of 2005 to the second quarter of 2009. (5) Within-effect penal regressions. County-specific intercept is not included in the fitted value of unemployment rates that will be used in the second stage.