Web Appendix to Industry Evidence on the Effects of Government Spending

By Christopher J. Nekarda and Valerie A. Ramey*

Input-Output Data

We use data from the Bureau of Economic Analysis (BEA)'s benchmark input-output (IO) tables to construct measure of industry-level government spending. Benchmark input-output (IO) tables based on Standard Industrial Classification (SIC) codes are available for 1963, 1967, 1972, 1977, 1982, 1987, 1992. Starting in 1997, the IO tables moved from a SIC-based classification to one based on the North American industrical classification system (NAICS). We do not use the NAICS–based IO tables because we fear merging industries based on NAICS-SIC correspondences may be fraught with additional error.

The IO tables are available on the U.S. Census Bureau's web site.[1] The table below lists the file names for the transactions and total requirements benchmark IO tables.

Year	*Source file*
Transactions	
1963	`1963 Transactions 367-level Data.txt`
1967	`1967 Transactions 367-level Data.txt`
1972	`1972 Transactions 367-level Data.txt`
1977	`1977 Transactions 366-detail Data.txt`
1982	`82-6DT.DAT`
1987	`TBL2-87.DAT`
1992	`SICUSE.TXT`
Total requirements	
1963	`1963 Transactions 367-level Data.txt`
1967	`1967 Transactions 367-level Data.txt`
1972	`1972 Total Requirements 365-level Data.txt`
1977	`1977 Total Req Coeff 366-level Data.txt`
1982	`82CCTR.TXT`
1987	`TBL4-87.DAT`
1992	`CXCTR.TXT`

* Nekarda: Board of Governors of the Federal Reserve System, 20th St & C St NW, Mailstop 80, Washington, DC, 20551, christopher.j.nekarda@frb.gov. Ramey: University of California, San Diego and NBER, 9500 Gilman Dr., La Jolla, CA 92093-0508, vramey@ucsd.edu.

[1] `http://www.bea.gov/industry/io_benchmark.htm`.

Except for 1963, the IO data are available at a 6-digit level (537 industries); in 1963, the data are available at 4-digit level (367 industries). All calculations are performed at the most disaggregated level available.

Let S_{ijt} be the value of inputs produced by industry i shipped to industry j in year t, measured in producers' prices. Direct government demand for industry i is the value of inputs from industry i used by the federal government ($j = g$):

$$G_{it}^{d} = S_{igt}. \tag{A.1}$$

Although the IO tables distinguish between defense and nondefense federal purchases, we take the sum of both categories. The table below lists the IO codes for the government for each benchmark IO table.

	Industry code	
Year	*Defense*	*Nondefense*
1963	9710	9720
1967	971000	972000
1972	960000	970000
1977	960000	970000
1982	960000	970000
1987	960000	970000
1992	9600I0, 9600C0	9700I0, 9700C0

Indirect government demand is calculated using commodity-by-commodity unit input requirement coefficients. Let r_{ijt} be the commodity i output required per dollar of each commodity j delivered to final demand in year t. The indirect government demand for industry i's output is the direct government purchases from industry j times the unit input requirement of industry i for industry j's output:

$$G_{it}^{n} = \sum_{j=1}^{J_t} G_{jt}^{d} \times r_{ijt}. \tag{A.2}$$

Total government demand for industry i in year t is the sum of direct and indirect demand:

$$G_{it} = G_{it}^{d} + G_{it}^{n}. \tag{A.3}$$

After calculating direct and indirect government shipments, the IO data are aggregated (see below) to merge with the Manufacturing Industry Database (MID) data.

Manufacturing Industry Database

Data on manufacturing industries comes from the National Bureau of Economic Research–Center for Economic Studies (NBER-CES) MID database.[2] The MID database contains annual data on 459 manufacturing industries from 1958 to 2005. The data are compiled from the Annual Survey of Manufacturers and the Census of Manufactures and adjust for changes in industry definitions over time. We use the version based on the 1987 SIC codes.

We use MID measures of gross shipments; employment, annual hours worked, and the wage bill for production and nonproduction workers; total capital; plant, equipment, investment, materials usage, and energy usage. The MID also includes price indexes for capital, investment, materials, and energy. We create real series from the nominal values by dividing by the appropriate price index. The production worker product wage is the production worker wage bill divided by production worker hours times the shipments deflator.

Total Hours

The database provides information on annual hours only for production workers. We created two measures of total hours using two extreme assumptions: nonproduction workers always work 1,960 hours per year and nonproduction workers always work as much as production workers. The constant-hours value is slightly less than the usual 2000 hours per year because it allows for vacations and holidays, which are not included in production worker hours measures. The results were very similar using both measures, so we only report the results using the conservative assumption that nonproduction workers' hours are constant.

Labor Share

The payroll data from the MID includes only wages and salaries; it does not include payments for benefits, such as Social Security and health insurance. Thus, labor share estimates from the MID are biased downward. Following Chang and Hong (2006), we use national income and product accounts (NIPA) data to compute the ratio of total compensation to wages and salaries for each 2-digit SIC manufacturing industry.[3] When the NIPA data migrate to the NAICS codes in 2001, we adjust the factor shares by the difference in the ratio in 2000.

We merge these factors to our 4-digit data and use them to magnify the payroll data to create more accurate labor shares.

Real Output

We construct real shipments by dividing nominal shipments by the shipments price deflator. However, because firms hold inventories, shipments are not necessarily equal to output. According to the standard inventory identity, real gross output, Y, is equal to real shipments, S, plus the change in real finished-goods and work-in-process inventories, I^F. The MID database reports

[2] Bartelsman, Becker and Gray (2000). The data are available at `http://www.nber.org/nberces/`.

[3] Compensation is from table 6.2; wage and salary accruals are from table 6.3.

only the total value of inventories, I, at the end of the year; it does not distinguish inventories by stage of process in the reported stocks.

Fortunately, we can back out the nominal change in materials inventories from other data in the MID. In particular, the measure of nominal value added, $\tilde{V}$, in the MID is defined as:

$$\tilde{V}_{it}^{\text{MID}} = \tilde{S}_{it} - \tilde{M}_{it} + \Delta\tilde{I}_{it}^{F}, \tag{A.4}$$

where $\tilde{M}$ is nominal materials cost.

Since total inventories is the sum of finished-goods, work-in-process, and materials inventories, I^M, the change in materials inventories can be inferred from the change in total inventories and the change in finished-goods and work-in-process inventories: $\Delta\tilde{I}_{it}^{M} = \Delta\tilde{I}_{it} - \Delta\tilde{I}_{it}^{F}$. Using this inventory relationship, we calculate real gross output as

$$Y_{it} \approxeq \frac{\tilde{S}_{it}}{P_{it}} + \left[\frac{\tilde{I}_{it}}{P_{it}} - \frac{\tilde{I}_{i(t-1)}}{P_{i(t-1)}}\right] - \frac{\Delta\tilde{I}_{it}^{M}}{P_{it}}, \tag{A.5}$$

where P is the price of output. This formulation for gross output is not exact because the last term, the change in real materials inventories, should be

$$\frac{\tilde{I}_{it}^{M}}{P_{it}} - \frac{\tilde{I}_{i(t-1)}^{M}}{P_{i(t-1)}}.$$

Unfortunately the MID does not have data on the stock of materials inventories at each point in time necessary. As a result, our measure of gross real output in equation A.5 understates production by

$$\frac{\tilde{I}_{i(t-1)}^{M}}{P_{i(t-1)}} \times \frac{P_{it} - P_{i(t-1)}}{P_{it}}, \tag{A.6}$$

which is the product of the real initial stock of materials inventories (valued at output prices) and the rate of inflation of output prices. According to BEA estimates of inventories and sales in manufacturing, the real stock of materials inventories is about 50 percent of monthly sales, or about 4 percent of annual sales. Even if annual inflation is as high as 10 percent, the bias would only be −0.4 percent.

Marginal Markup

This section describes how we estimate Nekarda and Ramey's (2010) marginal-average wage adjustment factor for 2-digit SIC manufacturing industries. The adjustment factor (their equation 7) is the ratio of the marginal wage to the average wage:

$$\frac{W_{\text{M}}}{W_{\text{A}}} = \frac{1 + \rho\theta\left(\frac{\mathrm{d}v}{\mathrm{d}h}\right)}{1 + \rho\theta\left(\frac{v}{h}\right)},$$

where here, h is average hours per worker, v is average overtime hours per worker, ρ is the premium for overtime hours, and θ is the fraction of overtime hours that command a premium.

In the Current Employment Statistics (CES) data on production workers in manufacturing, overtime hours are defined as those hours that are paid an premium so $\theta = 1$ by definition. The Fair Labor Standards Act requires that employers pay a 50 percent premium for hours in excess of 40 per week for covered employees. As most workers who earned premium pay received a 50 percent premium, we set $\rho = 0.5$. Thus, to construct the ratio of marginal to average wages we require data on v/h and $\mathrm{d}v/\mathrm{d}h$.

Unfortunately, the MID does not contain information on overtime hours. To obtain estimates of overtime hours, we first regress v on h for each 2-digit industry in the CES data:

$$(v/h)_{it} = \xi_{i0} + \xi_{i1} h_{it} + \omega_{it}. \tag{A.7}$$

The coefficients ξ_{i0} and ξ_{i1}, reported in table A.1, are then used with the annual hours data at the 4-digit level in the MID to construct overtime hours. The value of v/h is then directly calculated.

As in Bils (1987) and Nekarda and Ramey (2010), we estimate $\mathrm{d}v/\mathrm{d}h$ using the parametric specification

$$\begin{aligned}\Delta v_{it} = &\Big\{b_{i0} + b_{i1} t + b_2 t^2 + b_3 t^3 + c_1 \left[h_{i(t-1)} - 40\right] + c_2 \left[h_{i(t-1)} - 40\right]^2 \\ &+ c_3 \left[h_{i(t-1)} - 40\right]^3\Big\}\Delta h_{it} + a_{i0} + a_1 t + a_2 t^2 + a_3 t^3 \\ &+ d_{i1} \ln\left[N_{it}/N_{i(t-1)}\right] + d_{i2} \Delta \ln\left[N_{it}/N_{i(t-1)}\right] + e_{it},\end{aligned} \tag{A.8}$$

where all parameters listed as a function of i are allowed to differ across industries.[4] We estimate this equation on a panel of 2-digit industries using quarterly data. All hours and employment data are for production and nonsupervisory workers. We seasonally adjust monthly data for each industry, in the process removing outlier observations from holidays, strikes, and bad weather, and then take a quarterly average.

[4] See Nekarda and Ramey (2010) for details.

TABLE A.1—REGRESSION OF RATIO OF OVERTIME HOURS TO AVERAGE HOURS ON AVERAGE HOURS

	Constant		*Average hours*		
SIC	$\widehat{\delta}_0$	*Standard error*	$\widehat{\delta}_1$	*Standard error*	R^2
20	−0.377***	0.057	0.012***	0.001	0.297
21	−0.588***	0.056	0.017***	0.001	0.439
22	−0.464***	0.023	0.014***	0.001	0.781
23	−0.362***	0.012	0.011***	0.000	0.864
24	−0.529***	0.031	0.015***	0.001	0.702
25	−0.428***	0.023	0.013***	0.001	0.730
26	−0.304***	0.043	0.010***	0.001	0.363
27	−0.380***	0.030	0.012***	0.001	0.584
28	−1.018***	0.045	0.026***	0.001	0.782
29	−0.698***	0.080	0.019***	0.002	0.372
30	−0.620***	0.031	0.017***	0.001	0.760
31	−0.308***	0.023	0.009***	0.001	0.584
32	−0.634***	0.023	0.018***	0.001	0.856
33	−0.655***	0.014	0.018***	0.000	0.948
34	−0.721***	0.026	0.020***	0.001	0.853
35	−0.629***	0.030	0.017***	0.001	0.776
36	−0.625***	0.032	0.017***	0.001	0.732
37	−0.612***	0.025	0.017***	0.001	0.832
38	−0.683***	0.030	0.018***	0.001	0.780
39	−0.482***	0.033	0.014***	0.001	0.622

Source: Author's regressions using quarterly CES data.

Notes: Regression of $(v/h)_{it} = \xi_{i0} + \xi_{i1} h_{it} + \omega_{it}$ separately for each industry i. Sample contains 20 2-digit SIC industries over 1960:1–2002:4 (172 observations per industry). *** indicates significance at 1-percent, ** at 5-percent, and * at 10-percent level.

IO-SIC Correspondence

We develop a correspondence between the 6-digit IO code–based input-output data and the 4-digit SIC code–based Manufacturing Industry Database (MID) data. The ultimate correspondence is between 4-digit IO and 4-digit SIC codes, but aggregation is required in both data sets to achieve a one-to-one correspondence. Aggregation is preformed at the most detailed level possible. The merged database contains 272 industries at the 4-digit SIC level. The correspondences discussed below are included as `.csv` files in the supplementary materials.

Combined IO Codes

To create a complete correspondence, we occasionally need to assign a 6-digit IO code to a different 4-digit IO code for aggregation. For example, in the 1967 IO table, IO codes 170200 ("felt goods, n.e.c."), 170300 ("lace goods"), 170400 ("paddings & upholstery filling"), and so forth are assigned to 4-digit IO code 1710 ("textile goods, n.e.c.") for final aggregation. This is required because no such industries existed at the 4-digit SIC level.

The files `combined-io-`*`YYYY`*`.csv` lists which 6-digit IO codes are assigned to which 4-digit IO codes before aggregating to the 4-digit IO level for each year. This assignment changes for each benchmark IO table. We aggregate the nominal value of shipments by summing the component sectors.

Combined SIC Codes

Some SIC codes also require combination. This aggregation occurs mostly at the 3-digit SIC level, which roughly corresponds to the 4-digit IO level, but it is ultimately tailored to preserve the best correspondence with the IO data.

The real quantities were defined at the industry level as the nominal quantities divided by the relevant price index. Because the price indices in this data base are fixed-weight indices, it is possible to sum the real quantities. We then summed nominal and real quantities for the combined industries and used their ratios to construct price indices.

The file `combined-sic.csv` lists which 4-digit SIC codes are combined into a single sector. The combined value of the sectors in the source column is assigned to the SIC code listed in the destination column.

Correspondence between IO and SIC Codes

The file `io-sic-correspond.csv` lists the complete correspondence between 4-digit IO codes and 4-digit SIC codes. The codes used in this table reflect the assignments and combinations described in the previous two sections.

Other Data Sources

Item	*Source*
Average weekly hours and average weekly overtime hours	BLS Employment, Hours, and Earnings database: `ftp.bls.gov/pub/time.series/ee/`
Four-firm concentration rate	U.S. Census Bureau: `www.census.gov/epcd/www/concentration92-47.xls`
Unionization rate	Abowd (1990)
Nominal federal spending	NIPA table `1.1.5`, lines `22` and `23`
Implicit price deflators	NIPA table `1.1.9`, lines `22` and `23`

REFERENCES

Abowd, John M. 1990. "The NBER Immigration, Trade, and Labor Markets Data Files." National Bureau of Economic Research Working Paper 3351.

Bartelsman, Eric J., Randy A. Becker, and Wayne B. Gray. 2000. "NBER-CES Manufacturing Industry Database." National Bureal of Economic Research.

Bils, Mark. 1987. "The Cyclical Behavior of Marginal Cost and Price." *American Economic Review*, 77(5): 838–55.

Chang, Yongsung, and Jay H. Hong. 2006. "Do Technological Improvements in the Manufacturing Sector Raise or Lower Employment?" *American Economic Review*, 96(1): 352–68.

Nekarda, Christopher J., and Valerie A. Ramey. 2010. "The Cyclical Behavior of the Price-Cost Markup." University of California, San Diego Unpublished paper.