FEDERAL RESERVE BANK OF SAN FRANCISCO

WORKING PAPER SERIES

A State Level Database for the Manufacturing Sector: Construction and Sources

Robert S. Chirinko
Emory University and CESifo

Daniel J. Wilson
Federal Reserve Bank of San Francisco

October 2, 2009

Working Paper 2009-21
http://www.frbsf.org/publications/economics/papers/2009/wp09-21bk.pdf

A State Level Database For The Manufacturing Sector:

Construction And Sources

Robert S. Chirinko
Emory University and CESifo

and

Daniel J. Wilson*
Federal Reserve Bank of San Francisco

October 2, 2009

* The authors gratefully acknowledge the helpful comments and assistance offered by Jane Gravelle, Sharon Kozicki, and Huntley Schaller and the expert research assistance provided by Ann Lucas and Charles Notzon. Robert Chirinko acknowledges critical financial support provided by the Federal Reserve Bank of San Francisco. All errors and omissions are the sole responsibility of the authors, and the conclusions do not necessarily reflect the views of the organizations with which they are associated.

A State-Level Database For The Manufacturing Sector: Construction And Sources

Appendix

This document describes the construction of and data sources for a state-level panel data set measuring output and factor use for the manufacturing sector. These data are a subset of a larger, comprehensive data set that we currently are constructing and hope to post on the FRBSF website in the near future. The comprehensive data set will cover the U.S. manufacturing sector and may be thought of as a state-level analog to other widely used productivity data sets such as the industry-level NBER Productivity Database or Dale Jorgenson's "KLEM" database or the country-level Penn World Tables, but with an added emphasis on adjusting prices for taxes. The selected variables currently available for public use are nominal and real gross output, nominal and real investment, and real capital stock. The data cover all fifty states and the period 1963 to 2006.

A State-Level Database For The Manufacturing Sector: Construction And Sources

This document describes the construction of and data sources for selected variables used in Chirinko and Wilson (2008). These data are a subset of a larger, comprehensive data set that we currently are constructing and hope to post on the FRBSF website in the near future. The comprehensive data set will cover the U.S. manufacturing sector and may be thought of as a state-level analog to other widely used productivity data sets such as the industry-level NBER Productivity Database or Dale Jorgenson's "KLEM" database or the country-level Penn World Tables, but with an added emphasis on adjusting prices for taxes. The selected variables currently available for public use are nominal and real gross output, nominal and real investment, and real capital stock. The data cover all fifty states and the period 1963 to 2006.[1]

The state data described in this document measure economic activity in the manufacturing sector. The primary raw source data for the state-level totals of output, investment, labor and establishments counts is the Annual Survey of Manufacturers (ASM) conducted by the U.S. Census Bureau. State-level totals (which the Census Bureau refers to as "AS-3" data) are reported in the yearly volumes of the ASM publication. From 1994 onward, these data also can be found in the yearly ASM Geographic Area Statistics (ASM-GAS) publications. Hereafter, we will refer to the ASM data on state-level totals for all years as the ASM-GAS data. The ASM data are collected from a large, representative sample of manufacturing establishments with one or more paid employees. The 2004 ASM (Appendix B, p. B-1) defines the manufacturing sector as follows,

> "The Manufacturing sector comprises establishments engaged in the mechanical, physical, or chemical transformation of materials, substances, or components into new products. The assembling of component parts of manufactured products is considered manufacturing, except in cases where the activity is appropriately classified in Sector 23, Construction. Establishments in the manufacturing sector are often described as plants, factories, or mills and characteristically use power-driven machines and materials-handling equipment. However, establishments that transform materials or substances into new products by hand or in the worker's home and those engaged in selling to the general public products made on the same premises from which they are sold, such as bakeries, candy stores, and custom tailors, may also be included in this sector. Manufacturing establishments may process materials or may

[1] The data used in Chirinko and Wilson (2008) covered only 1982-2004. Our data set does not include the District of Columbia (D.C.) as the ASM-GAS data for D.C. appears to have been noticeably affected by the switch in 1997 from a SIC to NAICS basis for defining the manufacturing sector.

> contract with other establishments to process their materials for them. Both types of establishments are included in manufacturing."

The ASM manufacturing sector corresponds to NAICS sectors 31 to 33.

1. OUTPUT – $Y_{s,t}$

Output is measured by real value added, and it is defined as nominal value added divided by a price deflator,

$$Y_{s,t} = Y\$_{s,t} / P^{BT,Y}_{mfg,t},$$

where $Y\$_{s,t}$ is nominal value added output and $P^{BT,Y}_{mfg,t}$ is the price index for manufacturing output net of sales and excise taxes but before corporate income tax adjustments. Since the $P^{BT,Y}_{mfg,t}$ series is based on producer price indices, it measures average prices received by domestic producers (PPI). Our database presents $Y_{s,t}$ in billions of constant 2000 dollars.

The $Y\$_{s,t}$ series is obtained from ASM-GAS (e.g., in 2004, the data are published in Table 1, column F). Because the ASM was not conducted for the years 1979 to 1981, $Y\$_{s,t}$ is missing for all states for these years. In future versions of this data set, we intend to estimate these missing values, but for now we leave these data points as missing. The ASM-GAS data for this series were undisclosed for Minnesota for the years 1970 and 1971. We filled these two data points in via linear interpolation between Minnesota's value added data for 1969 and 1972. Our database presents $Y\$_{s,t}$ in billions of dollars.

The $P^{BT,Y}_{mfg,t}$ series is obtained from INDUSTRY, the table labeled "Chain-Type Price Indexes for Value Added by Industry," Line 12. Our database presents $P^{BT,Y}_{mfg,t}$ as an index number with a base year value in 2000 of 1.0.

2. CAPITAL -- $K_{s,t}$

Capital input is measured by the real (constant-cost) replacement value of equipment (excluding software) and structures, and this series is constructed from the following perpetual inventory formula,

$$K_{s,t} = K_{s,\tau}(1-\delta_{mfg,t})^{t-\tau} + I_{s,t} \qquad t = \tau+1,...,T \ ,$$

where $K_{s,\tau}$ is the initial value of the real capital stock (where the index τ represents the initial period), $\delta_{mfg,t}$ is the rate of economic depreciation (hence $(1-\delta_{mfg,t})$ is the survival rate), and $I_{s,t}$ is real total capital expenditure. The capital stock is dated end-of-period (EOP). Our database presents $K_{s,t}$ in billions of constant 2000 dollars. Each component determining the capital stock is discussed in the following subsections.

2.1. The Initial Value Of The Capital Stock -- $K_{s,\tau}$

The $K_{s,\tau}$ series is measured by the book value of the capital stock adjusted for inflation,

$$K_{s,\tau} = K_{s,\tau}^{BV} * \left(K_{mfg,\tau}^{CoC} / K_{mfg,\tau}^{HC} \right),$$

where $K_{s,\tau}^{BV}$ is the book value (historical-cost) of the capital stock for state s, $K_{mfg,\tau}^{CoC}$ is the constant-cost value of the capital stock for the manufacturing sector, and $K_{mfg,\tau}^{HC}$ is the historical-cost value of the capital stock for the manufacturing sector. All capital stock series are EOP. Inflation drives a wedge between book value capital stocks (based on the original purchase cost of investment) and real capital stocks useful in economic analyses. The $\left(K_{mfg,\tau}^{CoC} / K_{mfg,\tau}^{HC} \right)$ ratio provides an approximate adjustment for the inflation wedge based on national manufacturing industry data. Our database presents $K_{s,\tau}$ in billions of constant 2000 dollars.

We compute initial values of the real capital stock EOP for $\tau = 1962$ and $\tau = 1981$. Note we "re-initialize" the capital stock in 1981 (as opposed to simply using the perpetual inventory formula starting with the 1962 initial stock estimate) for two reasons. First, the 1962 initial stock is

estimated (as described below) rather than observed, since book value was not collected by the ASM prior prior to 1975, and so we do not want to rely too heavily on the 1962 initial value estimate. Second and more importantly, data on capital expenditures are missing for 1979 to 1981. Thus, the initial capital stock for 1981, based on book value data, likely is a better measure of the true capital stock in 1981 than a capital stock measure based in part on imputed investment data from 1979 to 1981.

A provisional estimate of $K_{s,1962}$, $K^{\#}_{s,1962}$, is estimated by solving backward using the perpetual inventory formula, beginning with the 1975 data on the book value of capital (adjusted for inflation), subtracting investment data from 1963 to 1975, and weighting these terms by survival rates,

$$K^{\#}_{s,1962} = K^{\#}_{s,1975}(1-\delta_{1975})^{-(1975-1962)} - \left(\sum_{j=0}^{(1975-1962-1)} (1-\delta)^{-(1975-1962-j)} I_{s,1975-j} \right)$$

$$K^{\#}_{s,1975} = K^{BV}_{s,1975} * \left(K^{CoC}_{mfg,1975} / K^{HC}_{mfg,1975} \right)$$

The first part of the first of the equations above starts with the 1975 book value of capital (adjusted for inflation) and adds back all of the 1962 capital stock that has depreciated between 1962 and 1975. The second part then subtracts all of the investments made from 1963 to 1975, after adding back to each year's investment the portion that has depreciated between 1962 and when the investment was made. In essence, this formula undoes all of the additions to and depreciation from the original capital stock of 1962 and subsequent investments from 1963 to 1975. Note we choose 1975 as the year from which to work backwards since it is the earliest year in which book value data are available from the ASM.

The final estimate of $K_{s,1962}$ is then obtained by rescaling the provisional state estimates by the national real capital stock total in 1962 from the BEA, $K^{CoC}_{mfg,1962}$. Specifically,

$$K_{s,1962} = K^{\#}_{s,1962} * \left(K^{CoC}_{mfg,1962} \Big/ \sum_{s=1}^{51} K^{\#}_{s,1962} \right).$$

A potential inconsistency exists in using the BEA data to rescale our provisional estimate based on ASM data. Software investment is included in the BEA data but excluded in the ASM data. During the early 1960's, the discrepancy introduced by software investment is negligible. In 1963, software investment was 1.3% of manufacturing investment (though software embedded or bundled in computers and other equipment is not reflected in this figure). The impact of software investment is likely less than this figure for two reasons. First, for the older vintages of investment entering the 1962 capital stock, their share is likely to be even smaller than 1.3%. Second, software depreciates more rapidly than other capital. It would seem safe to conclude that that the discrepancy owing to the different treatment of software investment is less than 1% of the 1962 capital stock.

The $K^{BV}_{s,\tau}$ series is obtained from ASM (e.g., in 1975, the data are published in Table 4, row 5). Our database presents $K^{BV}_{s,\tau}$ in billions of dollars. The $K^{CoC}_{mfg,\tau}$ series is the product of a quantity index and a base year value that converts the index into a real stock,

$$K^{CoC}_{mfg,\tau} = INDEXK^{CoC}_{mfg,\tau} * K^{CuC}_{mfg,t=2000},$$

where $INDEXK^{CoC}_{mfg,\tau}$ is the chain-type quantity index for the real capital stock and $K^{CuC}_{mfg,t=2000}$ is the base year value for the current-cost value of the capital stock for the manufacturing sector. Our database presents $K^{CoC}_{mfg,\tau}$ in millions of dollars. The $INDEXK^{CoC}_{mfg,\tau}$ is obtained from FIXED, Table 4.2, line 7, and this series is divided by 100. Our database presents $INDEXK^{CoC}_{mfg,\tau}$ as an index number with a base year value in 2000 of 1.0. The $K^{CuC}_{mfg,t=2000}$ datapoint is obtained from FIXED, Table 4.1, line 7. Our database presents $K^{CuC}_{mfg,t=2000}$ in millions of dollars.

The $K^{HC}_{mfg,\tau}$ series is obtained from FIXED, Table 4.3, line 7. Our database presents $K^{HC}_{mfg,\tau}$ in millions of dollars.

2.2. The Rate Of Economic Depreciation -- $\delta_{mfg,t}$

The $\delta_{mfg,t}$ series is measured by the flow of annual depreciation divided by the capital stock existing at the beginning of the year,

$$\delta_{mfg,t} = \frac{D^{CuC}_{mfg,t}}{K^{CuC}_{mfg,t-1}},$$

where $D^{CuC}_{mfg,t}$ is the current-cost flow of depreciation in manufacturing industries and $K^{CuC}_{mfg,t-1}$ is the current-cost capital stock in manufacturing industries. Our database presents $\delta_{mfg,t}$ in percentage points.

The $D^{CuC}_{mfg,t}$ series is obtained from FIXED, Table 4.4, line 7. Our database presents $D^{CuC}_{mfg,t}$ in millions of dollars.

The $K^{CuC}_{mfg,t-1}$ series is obtained from FIXED, Table 4.1, line 7. Our database presents $K^{CuC}_{mfg,t-1}$ in millions of dollars.

2.3. Real Total Capital Expenditure -- $I_{s,t}$

Real total capital expenditure is defined as nominal capital expenditures deflated by a price index,

$$I_{s,t} = \frac{I\$_{s,t}}{P^{I}_{mfg,t}},$$

$$I\$_{s,t} = I\$^{NEW}_{s,t} + I\$^{USED}_{s,t},$$

where $I\$_{s,t}$, $I\$^{NEW}_{s,t}$, and $I\$^{USED}_{s,t}$ are total, new, and used nominal capital expenditures, respectively, and $P^{I}_{mfg,t}$ is the price deflator for investment for the manufacturing sector. Our database presents $I_{s,t}$ in billions of constant 2000 dollars. The $I\$_{s,t}$ and $P^{I}_{mfg,t}$ series are discussed in the following subsections.

2.3.1. Total Nominal Capital Expenditure -- $I\$_{s,t}$

The $I\$_{s,t}$ series is obtained in three different ways each of which are based on the ASM-GAS and depend on disjoint time periods. (This mixture of direct and indirect estimates is forced upon us because of some anomalies in the ASM-GAS.) The series represents nominal expenditures on equipment (excluding software) and structures. Our database presents $I\$_{s,t}$ in billions of dollars.

For 1977, 1978, and 1982 to 2004, the series is obtained directly from ASM-GAS (e.g, in 2004, the data are published in Table 2, column I). For 1963 to 1976, the ASM-GAS only publishes data for $I\$^{NEW}_{s,t}$. For these years, $I\$_{s,t}$ is derived based on a state's mean ratio of $I\$^{NEW}_{s,t}$ to $I\$_{s,t}$,

$$I\$_{s,t} = I\$^{NEW}_{s,t} * MEAN_s\left\{I\$_{s,v} / I\$^{NEW}_{s,v}\right\}$$
$$t = 1963,...,1976$$
$$v = 1977,\ 1978,\ 1982,...,\ 2004.$$

where the $MEAN_s\{.\}$ is computed separately for each state and over all available observations represented by the index v.

For 1979 to 1981, the ASM was not conducted, and hence no ASM-GAS source data are available for $I\$_{s,t}$, $I\$^{NEW}_{s,t}$, nor $Y_{s,t}$. The missing investment data for these three years are estimated with the following three-step procedure. First, we rely on the availability of alternative output data from BEA for these three years and the workhorse of investment modeling, the accelerator model, to estimate the missing total capital expenditure data. Output is defined as real Gross State Product (GSP) for the manufacturing sector.[2] With these data and the available data for $I_{s,t}$, we estimate the following flexible accelerator model,

[2] For all intents and purposes, Gross State Product is conceptually identical to Gross Domestic Product, though small differences exist in some minor categories.

$$I_{s,t} / Y'_{s,t} = \alpha_s + \beta_{s,0}(\Delta Y'_{s,t} / Y'_{s,t-1}) + \beta_{s,1}(\Delta Y'_{s,t-1} / Y'_{s,t-2}) + \beta_{s,2}(\Delta Y'_{s,t-2} / Y'_{s,t-3}) + \varepsilon_{s,t} \quad , \quad t = 1977, 1978, 1982, ..., 2004$$

where α_s is a state-specific constant capturing state fixed effects, the β_s's are state-specific slope parameters, $\varepsilon_{s,t}$ is an error term, and $Y'_{s,t}$ is real manufacturing GSP. The $Y'_{s,t}$ series is nominal manufacturing GSP divided by a price deflator. Nominal manufacturing GSP is obtained from the BEA's Regional Economic Accounts (REA) data. (In 1997, the data are reported on both SIC and NAICS bases; we use the SIC figures.) The deflator is $P^{BT,Y}_{mfg,t}$ discussed in Section 1.

Second, we use the estimated parameters (represented by ^'s over the α and the β's), data for $Y'_{s,t}$ and $P^I_{s,t}$, and a transformed version of the above equation to generate a provisional estimate of $I\$_{s,t}$ ($I\$^{\#}_{s,t}$) for the missing nominal capital expenditure observations,

$$I\$^{\#}_{s,t} = \left[\hat{\alpha}_s + \hat{\beta}_{s,0}(\Delta Y'_{s,t} / Y'_{s,t-1}) + \hat{\beta}_{s,1}(\Delta Y'_{s,t-1} / Y'_{s,t-2}) + \hat{\beta}_{s,2}(\Delta Y'_{s,t-2} / Y'_{s,t-3}) \right] * Y'_{s,t} * P^I_{s,t} \quad . \quad t = 1979, 1980, 1981$$

Third, for each year (1979, 1980, 1981), we rescale states' nominal investment so that it equals the national total, $I\$^{ASM}_{mfg,t}$, which we estimate by applying the growth rate of the BEA's nominal private nonresidential fixed investment (net of software) for the manufacturing sector, $I\$_{mfg,t}$, to the previous year's value of national investment reported in the ASM. Specifically, we multiply each state's provisional estimate by the ratio of national manufacturing investment to the national sum of the provisional estimates,

$$I\$_{s,t} = I\$^{\#}_{s,t} * \left(I\$^{ASM}_{mfg,t} \Big/ \sum_{s=1}^{51} I\$^{\#}_{s,t} \right),$$
$$t = 1979, 1980, 1981$$

$$I\$^{ASM}_{mfg,1979} = \sum_{s=1}^{51} I\$_{s,1978} \left(\frac{I\$_{mfg,1979}}{I\$_{mfg,1978}} \right)$$
$$I\$^{ASM}_{mfg,1980} = I\$^{ASM}_{mfg,1979} \left(\frac{I\$_{mfg,1980}}{I\$_{mfg,1979}} \right).$$
$$I\$^{ASM}_{mfg,1981} = I\$^{ASM}_{mfg,1980} \left(\frac{I\$_{mfg,1981}}{I\$_{mfg,1980}} \right)$$

The $I\$_{mfg,t}$ series is obtained from FIXED, Table 4.7, line 7 less the sum of software investment over all manufacturing industries (NAICS sectors 31 to 33) from DETAILED, row 9

The ASM-GAS data for $I\$_{s,t}$ need to be adjusted for additional missing values and an error. The additional missing values occur because the ASM-GAS did not report data for Minnesota for the years 1970 and 1971. We use the relation between BEA data for the manufacturing sector and state data for Minnesota on investment expenditures to impute the missing values with the following relation,

$$I\$_{s=minnesota,t} = MEAN \left\{ \frac{I\$_{s=minnesota,v}}{I\$_{mfg,v}} \right\} * I\$_{mfg,t}$$
$$t = 1970, 1971$$
$$v = 1967, 1968, 1969, 1972, 1973, 1974$$

where $I\$_{mfg,t}$ is nominal capital expenditures on new and used capital by the manufacturing sectors defined above and the mean of the ratio is computed for three years before and after the missing values. The $I\$_{mfg,t}$ series was discussed previously in this subsection.

The error occurs for $I\$_{s=ohio,t=1996}$. In 1996, ASM-GAS shows a 400% jump in nominal total capital expenditures in Ohio from about $8 billion in 1995 to $40 billion in 1996 and then back down to $9 billion in 1997. This enormous jump can be traced to the motor vehicles sector ($35

billion), which suggests a huge capital investment – equal to 85% of the sector's national capital expenditures – for the building of an auto plant(s) in Ohio in 1996. We dismiss this number for three reasons. First, the magnitude of this investment is implausible. By comparison, DaimlerChrysler's jeep plant expansion in Toledo in 1998 was $1.2 billion of total investment over several years. Second, correspondence with experts on the Ohio manufacturing sector (including one at the Ohio Department of Economic Development) could not confirm any massive capital expenditure programs in 1996. Third, the 1996 value for national total capital expenditures reported in the ASM-GAS is inconsistent with and about $32 billion higher than a comparable figure reported in a separate ASM publication, Statistics for Industry Groups and Industries (ASM-SIGI). These two publications disagree on national capital expenditures only in 1996, suggesting an error is present. We thus conclude that $I\$_{s=ohio,t=1996}$ = $40 billion is erroneous.

We fill in the 1996 Ohio data point by simply taking national manufacturing capital expenditures from the alternative ASM publication, ASM-SIGI, and subtracting the sum of capital expenditures from all other states.

2.3.2. Price Deflator For Investment -- $\mathbf{P^I_{mfg,t}}$

The price deflator for investment is constructed as an implicit deflator,

$$P^I_{mfg,t} = \frac{I\$_{mfg,t}}{I_{mfg,t}},$$

where $I\$_{mfg,t}$ and $I_{mfg,t}$ are nominal and real total capital expenditures, respectively, for the manufacturing sector. Our dataset presents $P^I_{mfg,t}$ as an index number with a base year value in 2000 of 1.0.

The $I\$_{mfg,t}$ series was discussed in the preceding subsection (Total Nominal Capital Expenditure).

The $I_{mfg,t}$ series is the product of a quantity index and a base year value that converts the index into real investment expenditures,

$$I_{mfg,t} = INDEXI_{mfg,t} * I\$_{mfg,t=2000},$$

where $INDEXI_{mfg,t}$ is the chain-type quantity index for real investment expenditures and $I\$_{mfg,t=2000}$ the base year value for current investment expenditures. Our database presents $I_{mfg,t}$ in billions of dollars. The $INDEXI_{mfg,t}$ is obtained from FIXED, Table 4.8, line 7, and this series is divided by 100. Our database presents $INDEXI_{mfg,t}$ as an index number with a base year value in 2000 of 1.0. The series containing the $I\$_{mfg,t=2000}$ datapoint was discussed in the preceding paragraph.

3. LEGEND/REFERENCES

ASM: CENSUS, *Annual Survey of Manufactures, Complete Volume* (Various Years). Complete volumes are available until 1993. In later years, the data are presented in separate pamphlets, which include ASM-GAS and ASM-SIGI mentioned below.

ASM-GAS: CENSUS, *Annual Survey of Manufacturers, Geographic Area Statistics* (Various Years). Publications for the years 1994 to 2004 (except 1997 and 2002) are available online. These data are published on an establishment basis. The data are obtained from electronic or paper documents depending on the time period: 2004 (Census website); 2003 to 1972 (CD's purchased from Census); 1971 to 1963 (paper copies). URL: http://www.census.gov/mcd/asm-as3.html.

ASM-SIGI: CENSUS, *Annual Survey of Manufacturers, Statistics for Industry Groups and Industries* (1996).
URL: http://www.census.gov/mcd/asm-as1.html.

BEA: Bureau of Economic Analysis, U.S. Department of Commerce.
URL: http://www.bea.gov.

BLS: Bureau of Labor Statistics, U.S. Department of Labor.
URL: http://www.bls.gov.

BOP: Beginning-Of-Period t.

CENSUS: Bureau of the Census, U.S. Department of Commerce.
URL: http://www.census.gov.

Chirinko, Robert S., and Daniel J. Wilson. 2008. "State Investment Tax Incentives: A Zero-Sum Game?" *Journal of Public Economics*, 92(12): 2362–2384.

DETAILED: BEA, Detailed Fixed Assets Tables, Nonresidential Investment, Historical-Cost.
URL: http://www.bea.gov/bea/dn/FA2004/Details/xls/detailnonres_inv1.xls

EOP: End-Of-Period t.

FIXED: BEA, *Standard Fixed Asset Tables.*
URL: http://www.bea.gov/bea/dn/FA2004/SelectTable.asp.

INDUSTRY: BEA, *Gross-Domestic-Product-by-Industry Accounts*. URL: http://www.bea.gov/bea/industry/gpotables.

PPI: BLS, *BLS Handbook of Methods*, Chapter 14 Producer Prices.
URL: http://www.bls.gov/opub/hom/pdf/homch14.pdf.

REA: BEA, *Regional Economic Accounts: Gross State Product*.
URL: http://www.bea.gov/bea/regional/gsp/default.cfm?series=SIC.