Collaborative Micro-Productivity Project

Federal Economic Statistics Advisory Committee
December 12, 2014

Lucia Foster[1], Cheryl Grim[1], Sabrina Wulff Pabilonia[2], **Jay Stewart**[2], **Zoltan Wolf**[3], and Cindy Zoghi[2]

[1] U.S. Census Bureau
[2] Bureau of Labor Statistics
[3] Westat

We thank John Haltiwanger for comments. Any opinions and conclusions expressed herein are those of the authors and do not necessarily represent the views of the Bureau of Labor Statistics or the U.S. Census Bureau. All results have been reviewed to ensure that no confidential information is disclosed.

United States Census Bureau | U.S. Department of Commerce
Economics and Statistics Administration
U.S. CENSUS BUREAU
census.gov

BLS
BUREAU OF LABOR STATISTICS
U.S. DEPARTMENT OF LABOR

Overview

- BLS produces the official productivity statistics for the U.S. using aggregate data; these data are critical for understanding the economy
- Information on the *within-industry* dispersion of productivity in the U.S. economy enhances our understanding of the rich productivity dynamics
- Census and BLS are collaborating to create new measures of the within-industry dispersion of productivity

Motivation

- Understanding Growth and Survival of Businesses in the U.S. Economy
  - Aggregate Measures of Productivity (BLS)
  - Micro-Level Measures of Productivity (Census)
  - Reallocation of Activity Across Businesses (Census)
- Collaborative Micro-Productivity Project (CMP) brings these pieces together to enhance our understanding of the economy

Background: Dispersion

- Researchers used Census micro-level data from the manufacturing sector to produce productivity measures
  - Result 1: Some plants are much more productive than others even within the same narrowly defined industry
    - Syverson (2004, RESTAT) found plants at the 90th percentile of the productivity distribution are nearly twice as productive as those at the 10th percentile
  - Result 2: Productivity dispersion varies across sectors, time, and countries

Background: Reallocation

- Low productivity plants contract and exit while high productivity plants survive and expand
  - Implies that aggregate productivity growth is driven in part by reallocation
- Plants that export are generally high productivity plants
- Open Questions
  - How do low productivity plants survive?
  - Does this dispersion reflect dispersion of fundamentals or distortions?
  - Given time variation in dispersion within sectors, does this imply that businesses are subject to time varying second moment shocks?

Background: Macro Dynamics

- These findings have changed the way economists think about aggregate productivity growth, labor market dynamics, international trade and globalization
- The CMP researchers are building the connection between the micro and the macro level statistics

Project Goal

- Produce public domain statistics on dispersion of productivity within narrowly defined sectors over time on a regular basis
- Specific objectives
  - Microdata files with input, output, and productivity measures to be made available in secure Research Data Centers
  - Joint Census and BLS public-use data product on the within-industry dispersion of productivity (labor and MFP) in manufacturing

Project Benefits (1)

- Reconcile differences between published aggregate statistics and their counterparts constructed by aggregating microdata
- Statistics on productivity distribution could potentially be integrated with other public domain products on firm dynamics
  - Management practices
  - Entrepreneurship and high growth firms
  - Innovation

Project Benefits (2)

- Enhance use of existing collected data
  - Pilot a new way of tabulating existing Census data to provide valuable new information to our data user community
- If successful, produce similar statistics for sectors other than manufacturing

The BLS Productivity Program

- Productivity Statistics (annual growth rates and indexes)
  - Labor productivity
  - Multifactor productivity
  - KLEMS multifactor productivity
    - Industry (3-digit)
    - Detailed industry (4-digit)

KLEMS MFP (8/21/2014)

	Manufacturing	
	1987-2012	2007-2012
Productivity		
Multifactor Productivity	1.2	0.4
Output per hour of all persons	3.3	1.5
Output per unit of capital services	-0.5	-2.6
Sectoral Output	1.8	-1.2
Inputs		
Combined Inputs	0.6	-1.6
Labor hours	-1.4	-2.7
Capital services	2.3	1.5
Energy	-0.2	-6.0
Materials	1.6	-2.7
Purchased business services	1.2	-1.1

Data on Capital

- Capital rental prices
- Productive capital stock (productivity capacity of existing capital stock)
- Gross investment
- Investment price deflator
- Wealth stock (financial value of existing capital stock based on remaining service lives)
- Depreciation of wealth stock

Summary

- OPT publishes estimates of labor productivity growth and MFP growth for major sectors and detailed industries
- OPT does not publish levels, although indexes are available
- Published estimates show
  - The contributions of different inputs to output growth
  - The contributions of different industries to aggregate productivity growth
- But no information below the industry level

Project Data

- Establishment-Level
  - Annual Survey of Manufactures (ASM)
  - Census of Manufactures (CM)
  - Longitudinal Business Database (LBD)
- Industry-Level
  - Data from BLS, BEA, NBER-CES Manufacturing Industry Database
    - Capital costs and depreciation rates
    - Deflators

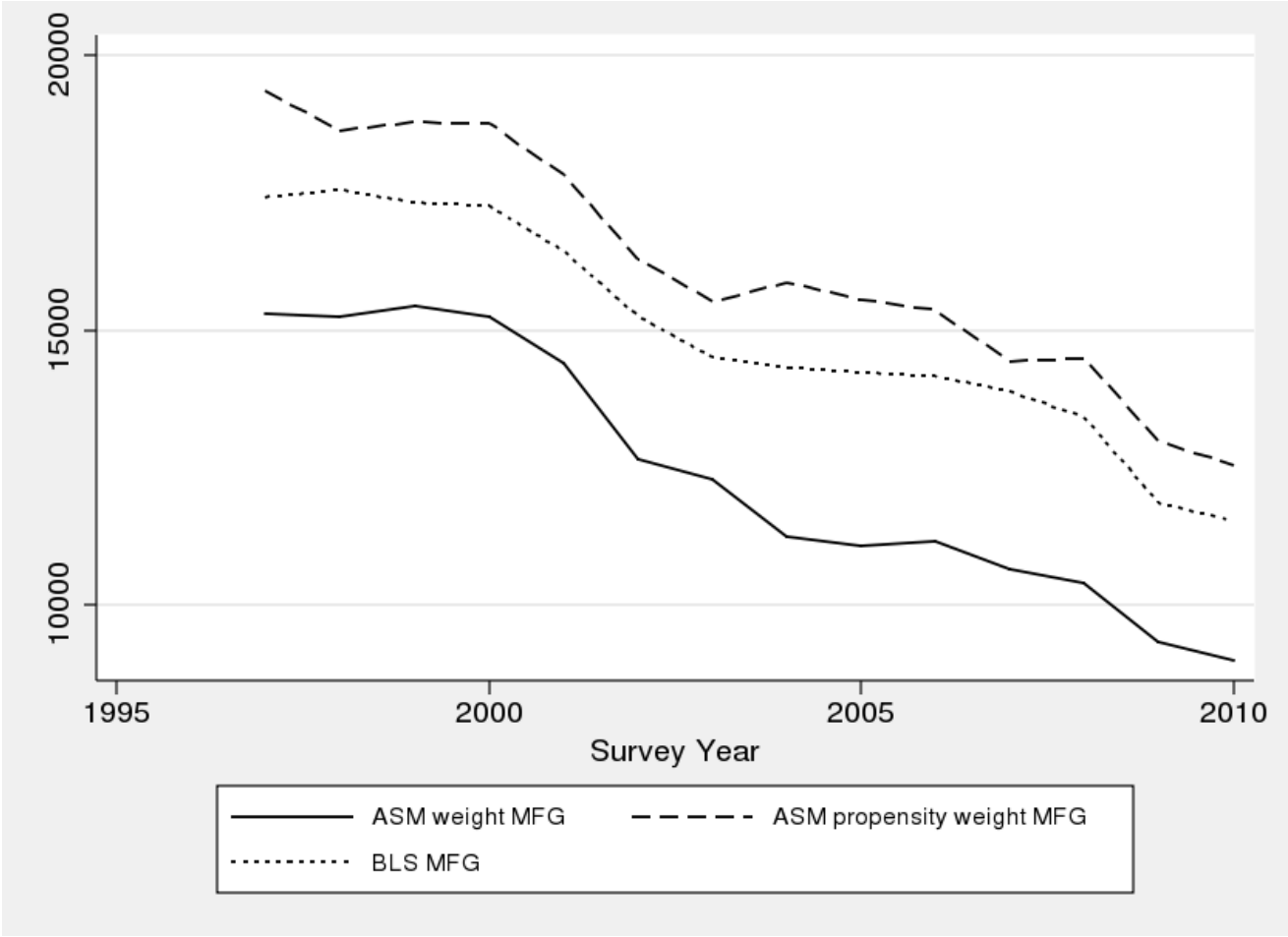
Establishment-level data

- ASM
  - Rotating panel between years ending in ‘4’ and ‘9’
  - Some establishments are sent a form while others are not
    - Certainty and non-certainty cases (sample weights)
    - Non-mail cases, based on payroll cutoff
- CM
  - Every 5 years in years ending in ‘2’ and ‘7’
  - ASM cases, large and medium sized establishments, small companies (short CM form)
  - Data for very small, single-establishment firms are imputed using administrative data
- LBD
  - Universe of business plants between 1976-2012
  - High quality longitudinal identifiers based on survey and administrative data

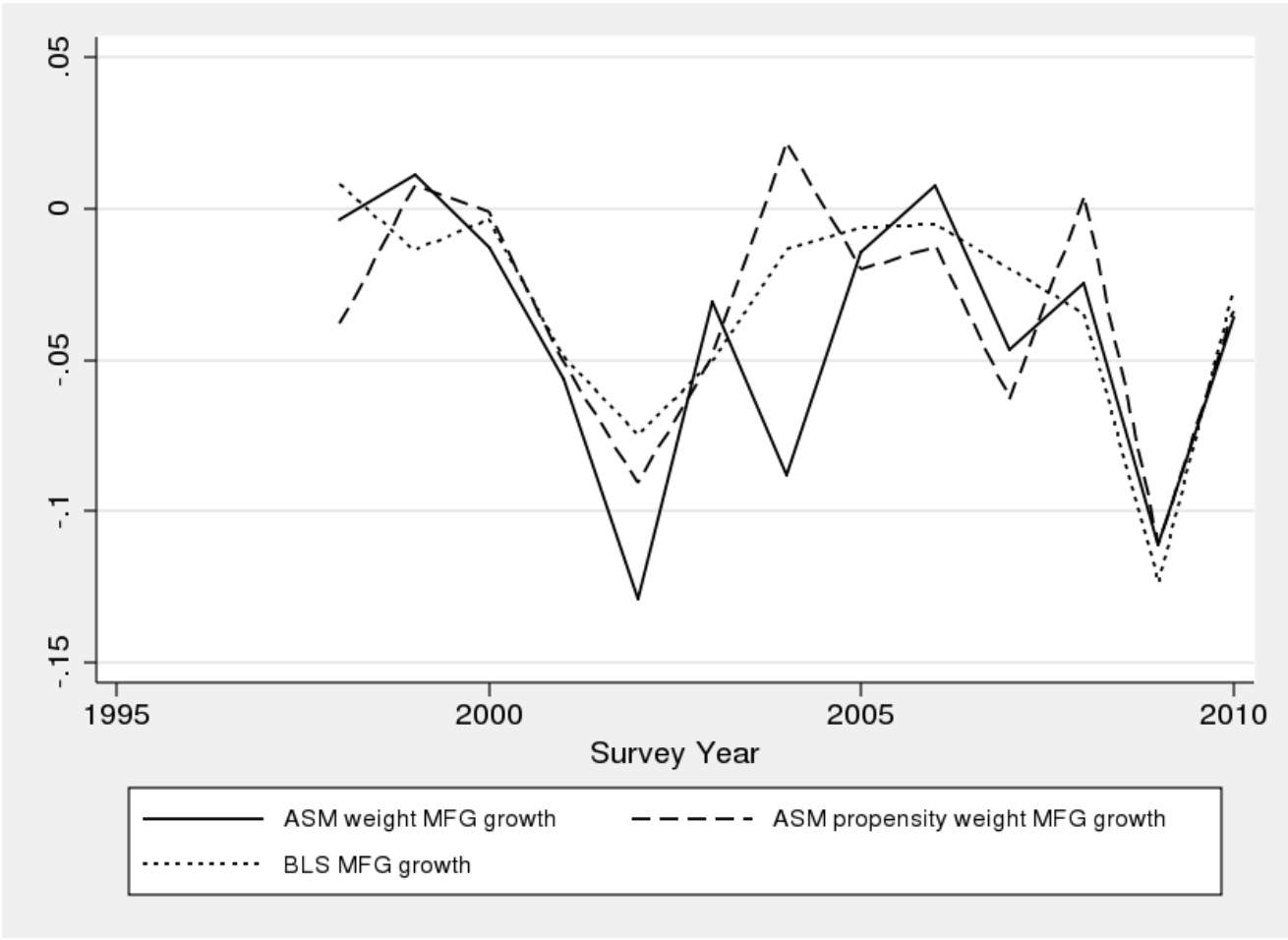
Industry-Level Comparisons

- Published industry-level BLS productivity measures are indices measuring changes over time within industries or broad sectors.
- Proposed dispersion measures are based on dispersion in **levels** of productivity within industries.
- What is the correlation between industry-level measures created from ASM/CM microdata and BLS measures?
  - First step, compare ASM/CM input and output measures aggregated to the industry level with BLS industry-level data

Employment Level



Employment Growth



How should we measure MFP?

- Multiple methods exist to calculate establishment-level MFP
  - Growth accounting (cost share) method
    - No econometric estimation required, flexible about exact shape of production technology, easy to implement
    - *But* requires strong assumptions about firm's optimizing behavior, at least in the long run.
  - Statistical methods
    - Flexible characterization of productivity, address endogeneity issue present in OLS, rely on establishment-specific variation to identify productivity shocks
    - *But* nonparametric approximations may exacerbate measurement error, some methods are computationally demanding, and small sample behavior is largely unexplored

Important Considerations for Measuring MFP

- Quality of measure
- Consistency with existing BLS measures, which are based on the growth accounting method
- Feasibility of calculation
- Impact of measure choice on resulting statistics

Growth Accounting MFP

MFP calculation involves many variables. Standard example (Cobb-Douglas production function):

$$\log(MFP_{et}) = \log(Q_{et}) - \beta_{it}^{K} * \log(K_{et}) - \beta_{it}^{L} * \log(L_{et}) - \beta_{it}^{E} * \log(E_{et}) - \beta_{it}^{M} * \log(M_{et})$$

where

e = establishment, *i* = industry, *t* = year
Q = output (deflated shipments revenue)

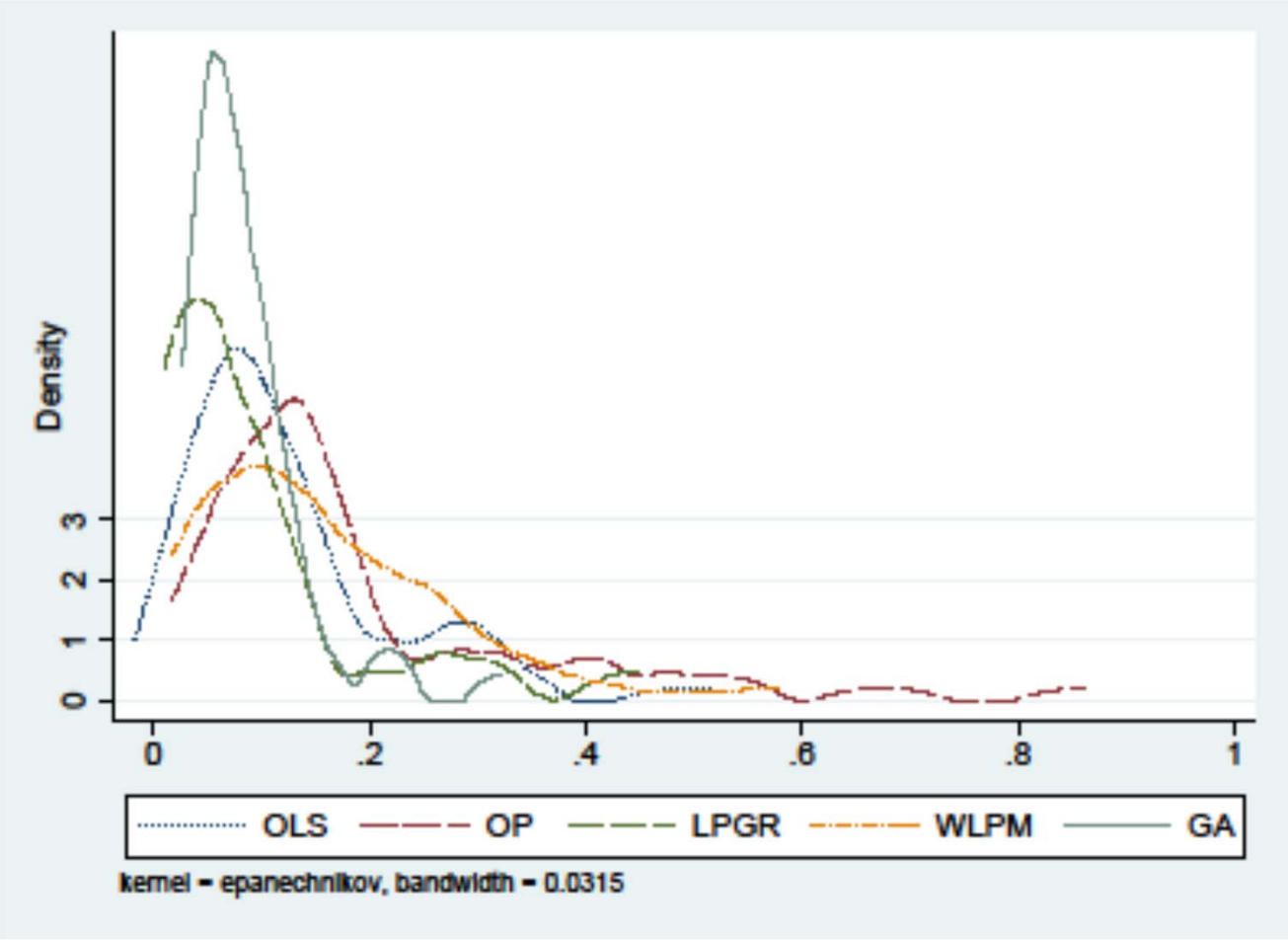
K = capital stock

L = labor (total hours worked)

E = deflated energy costs
M = deflated materials costs

β^{i} are factor elasticities for factor *i*

Capital Elasticity Distributions



Non-trivial differences in capital elasticity distributions across methods

Do elasticity differences matter for the general pattern of dispersion?

Method	Within-Industry Interquartile Range (IQR) of Establishment-Level Log MFP
Growth Accounting (GA)	0.24
Ordinary Least Squares (OLS)	0.26
Olley-Pakes (OP)	0.32
Levinsohn-Petrin Grid Search (LPGR)	0.29
Wooldridge Levinsohn-Petrin, Materials Proxy (WLPM)	0.40

Source: Foster, Grim, Haltiwanger, and Wolf (2014)
Note: Sample of 50 largest 4-digit industries

Differences in dispersion are nontrivial, but (for the most part) show consistently large differences in productivity across establishments.

Next Steps

- Obtain feedback from the research community
  - MFP calculation method
  - Imputation
  - Weighting
  - Disclosure avoidance
  - Desired statistics
- Create a beta version of the new data product

Questions for FESAC

- What data products would you like to see?
- Do you have suggestions on how to deal with some of our technical issues? Specifically:
  - The method for calculating MFP
  - Weighting and imputation

Background Slides

The remainder of the slides are for background purposes and provide more detail on the following areas:

1. Challenges
2. BLS productivity program
3. Project microdata

CHALLENGES

MFP Calculation Method

- Calculation of factor elasticities?
- Is log linear (consistent with a Cobb Douglas production function) sufficient?
- Limitations of data for outputs and inputs
  - No plant level prices for outputs and inputs
  - Capital stocks computed via perpetual inventory method. Challenges:
    - ASM panel rotation
    - Book values used to initialize
  - Non-production worker hours
  - Purchased services

Imputation

- Item level imputation rates are significant
  - Even for core variables such as total value of shipments, labor inputs, materials (the minimum needed to compute value added per worker)
- White, Reiter, and Petrin (2012) use the CART method to show that dispersion in MFP appears to be larger if imputation is taken into account
  - Their results are for a select number of products with physical quantity data in the Census of Manufactures
- We are considering alternatives including inverse propensity score weighting (IPW) for non-imputed data for addressing this issue (see, e.g., Wooldridge, 2002)
  - Preliminary results suggest using IPW for non-imputed data also yields higher measured dispersion within industries than dispersion measures computed from all data

Weighting

- Published statistics are not simply weighted sample totals
  - Adjustments for non-mail universe
- An alternative to using sampling probability weights is to use IPW for this purpose as well
  - Business Register/LBD provides universe from which ASM/CM is drawn

Disclosure Avoidance

- Plan is to release public domain statistics at the same level of detail as published ASM
  - In principle, this implies that disclosure avoidance should be able to take the same approach as ASM for measures such as the within industry standard deviation
- But given nature of data, robust dispersion measures (e.g., inter-quartile range) may be preferable

MORE DETAILS ON THE BLS PRODUCTIVITY PROGRAM

The BLS Productivity Program

- Productivity Statistics (annual growth rates)
  - Labor productivity
  - Multifactor productivity
  - KLEMS multifactor productivity

Labor Productivity News Release

- Nonfarm business sector
- Business sector
- Manufacturing
  - Durable
  - Nondurable
- Nonfinancial corporate sector

LP: Revised second-quarter 2014 measures (9/4/2014)

	Business	
	Q to Q	Y to Y
Productivity	2.3	0.9
Output	5.1	3.0
Hours	2.8	2.1
Hourly compensation	2.1	2.7
Real hourly compensation	-0.9	0.7
Unit labor costs.	-0.2	1.8

MFP News Releases

- MFP (capital and labor)
  - Private nonfarm business
  - Private business
- KLEMS (capital, labor, energy, materials, services)
  - Manufacturing
    - Durable
    - Nondurable
  - Industry (3-digit)
  - Detailed industry (4-digit)

MFP: Outputs, and Inputs (7/9/2014)

	Private nonfarm business	
	2000-2007	2007-2013
Productivity		
Multifactor Productivity	1.4	0.6
Output per hour of all persons	2.7	1.7
Output per unit of capital services	-0.5	-0.3
Output	2.8	1.0
Inputs		
Combined inputs	1.4	0.4
Labor input	0.4	-0.1
Hours	0.1	-0.6
Labor composition	0.3	0.5
Capital services	3.3	1.4

KLEMS MFP (8/21/2014)

	Manufacturing	
	1987-2012	2007-2012
Productivity		
Multifactor Productivity	1.2	0.4
Output per hour of all persons	3.3	1.5
Output per unit of capital services	-0.5	-2.6
Sectoral Output	1.8	-1.2
Inputs		
Combined Inputs	0.6	-1.6
Labor hours	-1.4	-2.7
Capital services	2.3	1.5
Energy	-0.2	-6.0
Materials	1.6	-2.7
Purchased business services	1.2	-1.1

Data on Capital

- 90 asset types
- Information processing capital (communication, computers, other)
- Intellectual property (software, R&D, artistic originals)
- Major sectors and NIPA-level industries

Data on Capital

- Capital rental prices
- Capital income
- Productive capital stock (productivity capacity of existing capital stock)
- Capital composition (ratio of capital input to productive stock – similar to labor composition)
- Gross investment
- Investment price deflator
- Depreciation of wealth stock
- Wealth stock (financial value of existing capital stock based on remaining service lives)

Data on Labor

- Hours worked for detailed industries
  - Average weekly hours
  - Employment
  - Total hours
- Constructed from
  - Current Employment Statistics survey
  - Current Population Survey
  - National Compensation Survey

Other Data

- Hourly compensation
- Unit labor costs
- Labor share
- Capital measures
- Cost shares
- Average weekly hours
- Employment

Summary of BLS Productivity Program

- OPT publishes estimates of labor productivity growth and MFP growth for major sectors and detailed industries
- OPT does not publish levels, although indexes are available
- Published estimates show
  - The contributions of different inputs to output growth
  - The contributions of different industries to aggregate productivity growth
- But no information below the industry level

MORE DETAILS ON THE PROJECT MICRODATA

Annual Survey of Manufactures

- The ASM is a series of 5-year panels
- New panels start in years ending in ‘4’ and ‘9’
- Establishments are split into two strata
  - Mail
    - Certainty cases – all receive a form
    - Non-certainty cases – sampled with probability based on a measure of size where larger establishments are more likely to be sent a form
  - Nonmail
    - Classified into nonmail based on a payroll cutoff that varies by industry
    - Consists of small and medium-sized, single establishment companies
- Sample weights are the inverse of selection probabilities
- Published statistics are not simply weighted micro data
  - Adjustments for non-mail universe

Census of Manufactures

- Every 5 years in years ending in ‘2’ and ‘7’
- Not all establishments receive a form
  - Sent a form
    - ASM establishments (ASM form; ~18% of all manufacturing establishments in 2007)
    - Large and medium establishments (Regular CM form; ~30% of all manufacturing establishments in 2007)
    - Small, single-establishment companies (Short CM form without detailed product and material questions; ~13% of all manufacturing establishments in 2007)
  - Not sent a form
    (~39% of all manufacturing establishments in 2007)
    - Very small, single-establishment companies
    - Annual payroll cutoff that varies by 6-digit NAICS
    - Data is imputed from administrative data

Longitudinal Business Database

- Census of business establishments and firms in the U.S. with paid employees
- Covers all industries and all U.S. states
- 1976-2012
- Establishment level
- Based on a combination of survey and administrative records
- Provides high quality longitudinal establishment and firm identifiers

Overview of MFP Estimation Methods for Microdata

Method	Description	estimator
OLS	Ordinary least squares	Least squares
OP	Olley and Pakes (1996)	Least squares, GMM, investment proxy
LPGR	Levinsohn and Petrin (2003)	Least squares, GMM, materials proxy
WLPM	Wooldridge (2009)	GMM, materials proxy
GA	Foster, Haltiwanger and Krizan (2001)	Industry-level cost-shares (non-statistical)

References

- Dean, Edwin R. and Michael J. Harper. 2001. “The BLS Productivity Measurement Program.” In *New Developments in Productivity Analysis*, ed. Charles R. Hulten, Edwin R. Dean, and Michael J. Harper, 55-84. Chicago and London: University of Chicago Press.
- Foster, Lucia, Cheryl Grim, John Haltiwanger, and Zoltan Wolf. 2014. “Macro and Micro Dynamics of Productivity: Is the Devil in the Details?” Unpublished manuscript.
- Foster, Lucia, John C. Haltiwanger, and Cornell J. Krizan. 2001. “Aggregate Productivity Growth. Lessons from Microeconomic Evidence.” In *New Developments in Productivity Analysis*, ed. Charles R. Hulten, Edwin R. Dean, and Michael J. Harper, 303-72. Chicago and London: University of Chicago Press.
- Harper, Michael J. 2007. “Technology and the Theory of Vintage Aggregation.” In *Hard-to-Measure Goods and Services: Essays in Honor of Zvi Griliches*, ed. Ernst E. Berndt and Charles R. Hulten, 99-120. Chicago and London: University of Chicago Press.
- Levinsohn, James and Amil Petrin. 2003. “Estimating Production Functions Using Inputs to Control for Unobservables.” *The Review of Economic Studies*, 70 (2): 317-41.
- Olley, Steven G. and Ariel Pakes. 1996. “The Dynamics of Productivity in the Telecommunications Equipment Industry." *Econometrica*, 64 (6): 1263-97.
- Syverson, Chad. 2011. “What Determines Productivity?” *Journal of Economic Literature*, 49(2): 326-65.
- White, T. Kirk, Jerome P. Reiter, and Amil Petrin. 2012. “Plant-level Productivity and Imputation of Missing Data in U.S. Census Manufacturing Data.” NBER Working Paper No. 17816.
- Wooldridge, Jeffrey M. 2009. “On Estimating Firm-level Production Functions Using Proxy Variables to Control for Unobservables.” *Economics Letters*, 104: 112-14.
- Wooldridge, Jeffrey M. 2002. “Inverse Probability Weighted M-Estimators for Sample Selection, Attrition, and Stratification.” Portugese Economic Journal, 1: 117-39.