State export data: origin of movement vs. origin of production

Cassey, Andrew

University of Minnesota

15 October 2006

State Export Data:
Origin of Movement vs. Origin of Production

Andrew J. Cassey*
Department of Economics
University of Minnesota

October 2006

Abstract

The Origin of Movement (OM) series is unique data documenting the destination of state exports. This data indicates the state an export begins its journey, not the production location (OP). Recent OM data has not been examined to determine if it represents OP. Here the collection, dissemination, and limitations of the OM data are described. Diagnostic tests asses how effectively the OM data represents OP. Results indicate the OM data are usable for OP, though there are idiosyncratic subsectors and states, and systematic differences distinguishing the OM from OP.

JEL classification: F14, R12, R14.

Keywords: international trade, exports, states, origin of movement, origin of production

---

*Thanks to my advisors: Sam Kortum and Tom Holmes. Thanks to Jaromir Nosal, Sasha Vostruknutov, Daisy Liu, Yoshi Kurokawa, and participants at the Workshop for Young Economists (Universidad de Guanajuato, Aug. 4–6th, 2006), the Missouri Valley Economics Association 43rd Annual Meeting (Minneapolis, MN, Oct. 26–28, 2006), and the Midwest International Trade Fall Meeting (Purdue University, Oct. 13-15, 2006) for comments. Partial funds for this project provided by Tom Holmes and the Department of Economics, University of Minnesota.

Correspondence: 1035 Heller Hall, 271 19th Ave S. Minneapolis, MN 55455; cassey@umn.edu
Department fax: 612 624 0209

1 Introduction

The Origin of Movement (OM) series is unique data on exports by U.S. states to foreign destinations. Compiled by the Foreign Trade Division of the U.S. Bureau of the Census, the OM series is a potentially useful data set since knowing both the location of the exporting state and the foreign destination is crucial for documenting facts about patterns of trade within the borders of the United States, patterns and characteristics of trading partners, and the effectiveness of state trade policy.[1] The OM data, however, is intended to indicate the state from which the export begins its journey abroad. It is not designed to represent the origin of production (OP) of exports, thus conceivably limiting its applicability.

Though frequently used by government agencies, policy institutions, regional researchers, and the business press, the OM series data has not received widespread attention from academics.[2] This lack of attention is due in part to (unverified) fears the OM data cannot represent OP. These fears are based on Census warnings, intuition, and a few eyebrow-raising observations, but not on any rigorous comparison or statistical test.

The following sections introduce the OM data to an academic audience and determine the substitutability of the OM data for OP using data from recent years. Section 2 describes the OM series and explains the details of how it is collected. Section 3 discusses the limitations of the OM data, explaining why the origin of movement is conceptually different from the origin of production of exports. Section 4 compares the OM data with another Census data set that does not contain destinations. This second data set is "Exports From Manufacturing Establishments" referred to as AR-1. Though also flawed, the best data on OP is the AR-1 data. Thus section 4 compares the OM and AR-1 data determining the overall substitutability of the data sets, as well as identifying any individual subsectors or states that are not substitutable. Results established here are in agreement with two previous studies using different technics and older data.

Although the comparison of the OM to AR-1 data reveals the OM data is, in general, of high

[1] The Census discontinued the Exporter Location (EL) series, begun in 1992, in December 2002. This data set is discussed further in appendix A. Also, there is a Canadian data set with exports of each province to each state in the U.S. but not each country in the world.

[2] Some of the peer-reviewed papers that do use the OM data are T. Smith (1990), P. Smith (1999), and Coughlin and Wall (2003).

enough quality to use as OP of exports, some individual states or subsectors differ between the two data sets. These differences are often systematic; there are industrial and geographic patterns. Section 5 documents these patterns and suggests modifications to the OM data that are desirable for many applications where OP is needed.

There are two important points before proceeding. First, a specific application or research questions is not specified here. Some research agendas call for using the OM data as OM. In this case there is no problem. Some agendas call for value-added of state exports rather than OP. This issue is not addressed but neither the OM nor AR-1 data try to measure value added. Other agendas call for origin of production of state exports, the focus here. Though no application of the data is specified here, any research involving the state of final production of exports benefits from the quality tests performed. The hope is this work stimulates research interest into state exports by destination by assuring the quality of the data. The second point is the quality of national export data is assumed. Thus there are no federal balance of payments, national trade deficit, or national trade program issues caused by the OM data differing from OP.

2 Description and Collection of the OM Data

The OM series data set contains current year export sales or free-alongside-ship (f.a.s.) costs[3] if not sold, by year, from fifty-four "states" (all 50 U.S. states and the District of Columbia, Puerto Rico, U.S. Virgin Islands, and "unknown") to 242 foreign destinations. Data begins in 1987 and runs through the present.[4] It is available quarterly or annually. Not only does the OM series contain states and destinations, it also provides data by industrial groups. Exports from each state to each destination are classified by thirty-three two digit Standard Industrial Classification (SIC) major groups for the years 1987 through 2000 or thirty-two three digit North American Industry Classification System (NAICS) subsectors for the years 1997 onward. Data by two, four, or six

[3]F.a.s. cost is the value of exports at port including the price of inland freight, and other charges but does not include loading fees or transit costs after port.

[4]In 2006, a sister series was created with the ZIP code of the origin of movement of an exported good. This series will have the same characteristics as the OM series except a finer level of geographic detail. In some sense, this new series continues the mission of the EL series corrected for the wild changes in data that lead to the cancelation of that series.

digit Harmonized System (HS) codes are available beginning in 1996 and are published monthly. Only data on physical commodities are collected; services are not. Table 13 in appendix C lists the NAICS and SIC subsectors and table 14 lists the destinations in the OM. The OM data are decomposed by mode of transportation: sea, air, and other.

The Census Bureau and Customs and Border Protection require exporters to provide information that becomes the OM data. The requirement affects any shipment of greater than $2500 in f.a.s. sales as well as any shipment of munitions or other specified goods. Until late 2005 an exporter completed a Shipper's Export Declaration (SED; included on page 47) collected at the port of exit or U.S. Post Office or an equivalent electronic format known as the Automated Export System (AES) filed before the export arrives at the port. Since late 2005 filing by AES is mandatory. Paper SEDs are no longer accepted, though for convenience, SED will henceforth refer to both the Shipper's Export Declaration and the AES. The SED has gone through multiple revisions over time. The only revision affecting the OM series data collection is the revision in 1985 adding the question on the origin of movement.

The SED defines an exporter as the seller of the commodity or the United States principal party in interest (USPPI) in the transaction. A shipment is defined as "All merchandize sent from one USPPI to one foreign consignee, to a single country of ultimate destination, on a single carrier, on the same day" (*Correct Way To Complete the Shipper's Export Declaration Form 7525-V*). Single carrier means each truck, railcar, ship, or airplane. The country of ultimate destination is the location of the receiving party of interest as determined by the Census's schedule C. Schedule C is more detailed than United Nations country codes. For example, schedule C contains a code for Heard and McDonald Islands, an uninhabited sub-arctic property administrated by Australia, whereas the United Nations lists Australia only. Schedule C is updated periodically, the most recent being November 2003. If the receiving party in interest redistributes or resells the export to another destination, the secondary destination is not recorded. Shipments may contain several commodities where each is defined as a single entry in the Census Bureau's schedule B classification on export commodities. Section 3 discusses how data collected by schedule B is disseminated as SIC or NAICS. Each schedule B commodity in a shipment must be declared separately. A re-export is indicated on the SED for any merchandise that enters the United States and is then exported

without modification. Finally, foreigners who export from the United States must complete an SED. Thus a car built by Toyota in Kentucky shipped to Mexico requires an SED.

An SED is not required on shipments in which there is no single commodity with sales exceeding $2500 unless that shipment is intended for Cuba, Iran, Iraq, Libya,[5] North Korea, Serbia, Sudan, and Syria, or if the commodity is on the *Commerce Control List*. An export valued under $2500, or any other exemptions, are indicated as such on the bill of lading, air waybill, or an equivalent loading document. An SED is not required for foreign shipments moving through the United States, although it is required for re-exports. Other exemptions described in *Correct Way to Complete the Shipper's Export Declaration Form 7525-V*.

Beginning in 1990, the SED is not required for exports to Canada unless the export falls under the *Commerce Control List*. Statistics Canada provides import data to the United States which is used for U.S. exports. Canadian regulations on imports are more stringent than United States requirements on exports. They require an import document with information equivalent to an SED be filed for a shipment with total value of $900 (Can.), even if no specific commodity is valued at $900. The Census takes the Canadian data and first converts values from Canadian to American dollars. The exchange rate is updated monthly. The Census then records Canadian imports by schedule B code for all items on a shipment whose total shipment value is over $900 (Can.). Individual commodities valued less than $2,500 (U.S.) are combined and recorded as a low value shipment and not reported in the OM data. The Census warns of comparability issues with Canadian export data pre- and post-1990.

The SED is the source of all official trade data of the United States. Exporters are required by law to accurately complete an SED. No information from the SED is released or shared with other agencies including the Internal Revenue Service. Therefore it seems unlikely for an exporter to lie.On the other hand, there are no incentives to accurately complete the SED either and estimates indicate that as many as 50% of paper SEDs have at least one mistake (Wysocki 1998). Electronic submissions are thought to contain far fewer errors suggesting newer OM data has fewer errors than older data.

Unlike other Census data, the OM series is not subject to disclosure concerns. A non-entry

[5]Libya was removed from this list of destinations on April 29, 2004.

indicates a value of less than $2,500 rather than a disclosure edit. The Census estimates the sum of all low value shipments, by country. Low value estimates are reported as their own schedule B number. They are not added to the classification of exports by subsector, and thus are not reported in the OM series.

State export data is collected from question 6 of the SED. This question instructs exporters to indicate the two letter U.S. Postal Service code indicating the state where the shipment begins its journey to the port of export or the identifying number of a free trade zone (FTZ). Data from FTZs are reported in the OM series as the state in which the zone is located. It is important to note the location of where the export begins its journey may differ from the production location of the export. This is due to several reasons. First, as directed by the SED, if in transit the shipment combines with similar commodities from the same USPPI, then the state of consolidation is indicated. If the shipment is multi-product, then the state of origin of movement for the single commodity of greatest value is indicated. Also, if the exporter is a wholesale or retail trader, then the location of the wholesale or retail establishment is the origin of movement. The location of a wholesaler or retailer need not be the same state as the producer. Finally, it is possible the location of origin of movement and production differ because the export is produced in one state and then sent to a warehouse in another before shipment abroad.

In the early years of the OM series, the origin of movement question frequently was unanswered on the SED; 25% and 20% of SEDs in 1987 and 1988 respectively. Beginning in 1988, the Census allowed the Massachusetts Institute for Social and Economic Research (MISER) to estimate the origin of movement of incomplete forms using an algorithm, causing some to refer to the OM (and EL) series as the "MISER data." In July 2004, MISER changed its name to the World Institute for Strategic Economic Research (WISER). Even with the MISER—now WISER—algorithm, exports without known origin are sizeable during the early 1990s. Due to the increased use of electronic filing, which has greatly reduced the number of incomplete forms, WISER has not made any adjustments to the raw data from the Census since 2000. Now WISER manages and sells the OM data.[6]

[6]Six years worth of state Export data by six digit HS codes since 2000 is available from Global Trade Information Services, Inc., `http://www.gtis.com`.

3 Limitations of the OM Data

The OM data's unique feature is both the origin of movement of an export and its destination are known. Therefore one can use the data to document export patterns by state, or export patterns by destination. That the origin of movement of an export is not necessarily the origin of production of the export limits the usefulness of the data. This is the primary limitation of the OM series and is discussed fully in section 3.1. It is not the only limitation however. Details of the full list of limitations are below, beginning with innate limitations and followed by limitations due to changes in data collection over time.

3.1 Innate Limitations

The OM series is limited in the information it provides from the SED. The OM series only reports export sales assigned to the state answered in question 6 on the SED, despite all of the other useful information obtained by the SED. For example, the OM does not indicate if the export is to a related party or not. Therefore one cannot use the data to learn about trade behavior of multinationals. Furthermore, since the exporter identification is not reported, there is no way of knowing the export behavior of small, medium, or large firms, or of multi-locational firms. For example, to test the hypothesis that state expenditures on export promotion activities matter for export sales, it is necessary to separate out multi-state firms since a trade contact made by the CEO whose headquarters is in Chicago, but whose factory and export depot is in Seattle, would credit the export promotion activities of Washington rather than Illinois. Similarly, the OM reveals nothing about the characteristics of exporting firms, such as their employment, or the proportion of firms that export directly. *A Profile of U.S. Exporting Companies* provides information about exports to related parties, the number of exporters, and export volume through wholesalers or directly from producers. This information is, however, not broken down by state, and so adds little when brought together with the OM data.

Another limitation of the data is the level of aggregation. The OM series data is for the subsector level: two out of four digit SIC or three out of six digit NAICS. The reason for this level of aggregation is problems of concordance between the export classification scheme used by the

Census and the classifications schemes for which the data is reported. The Census administers an export classification system known as schedule B: a list of approximately 8,000 entries. Schedule B is a commodity classification system rather than a production classification system such as SIC or NAICS. Schedule B groups similar goods together regardless of the inputs whereas NAICS or SIC distinguishes the same good if the production processes differ. If the export comes from a wholesaler, the exporter may not know the production process used. Therefore, there may be problems of assigning schedule B values to highly disaggregated NAICS or SIC. This mid-level of aggregation limits the researcher since all information about further levels of disaggregation is lost. The mid-level of aggregation has its benefits, though, as differences between the commodity classifications and production classifications at this level are negligible. A greater problem is that schedule B contains no codes corresponding to wholesale or retail since the value added of these firms does not substantially alter the physical object being exported. The sale value of the export is credited to the manufacturing, mining, or agriculture industries that primarily produce the physical object.

The OM series is used by state governments and regional researchers to determine the impact of trade on the state economy, the proportion of jobs related to exports, or the influence of state export promotion expenditures on the economy. Coughlin and Pollard (2000 and 2001), Kehoe and Ruhl (2004), or *Minnesota Quarterly Export Statistics*, are examples. Though used in this manner, the OM series is not intended to answer such questions. There are at least two reasons why the OM series may mislead when determining the effect of exports on a state's employment. First, as mentioned several times already, the OM series is not necessarily the OP of exports. Second, the OM series only reports direct exports. It says nothing about intermediate goods used in the production of the final export good. This is because the OM series reports export sales by state rather than value added of exports by state.

The following example illustrates the misleading results one can obtain from the OM data. Suppose Minnesota and Nebraska each produce agricultural products, NAICS 111, for export to Canada. All agricultural products in this example are exported by wholesalers. Minnesota grows \$95 worth of 111 which then goes through a Minnesotan wholesaler who adds nothing to the physical commodity, but does add \$5 of value added for its services. Final sales price is \$100. Now suppose

$47.50 of Nebraskan agricultural products are sold to a Nebraskan wholesaler who then directly exports it for $50 and $47.50 worth of agricultural exports are sold to a Minnesota wholesaler who directly exports it for $50. In this case, the OM series reports the value of agricultural exports from Minnesota in subsector 111 at $150 and the value of agricultural product exports from Nebraska in 111 at $50, for a 75%–25% share respectively. The OP indicates a 50%–50% share between Minnesota and Nebraska since $95 worth of agricultural products are produced in each state. Also note the $7.50 and $2.50 in value added from the wholesale industry is recorded as an agricultural product.

Suppose a researcher used the OM data generated by this example to make claims about the effectiveness of state export promotion expenditures on agricultural employment. Though incorrect, it appears Minnesota is much more successful than Nebraska in export driven agricultural employment.

Consolidation is not the only potential source of error in the data due to geography. The OM data is f.a.s. Presumably, there are greater transportation costs for interior states than for border states. This will tend to exaggerate the value of exports from interior states.

3.2 Time Series Limitations

Changes in data collection occur over the years that cause breaks in the OM time series. Though the way the Census collects the information for the OM series has not changed more than the shift from paper SEDs to electronic submission, other changes in data collection do affect the comparability between years. Major changes include modifications to the schedule B classifications of exports, the switch from collecting SEDs on Canadian exports to using Canadian import data in 1990, and the switch from SIC to NAICS.

Schedule B codes change frequently. Most of these changes are relatively small and involve one or two commodities at the 10 digit level, thus having no effect on aggregate data. In 1989, however, a major revision occurred as the Census matched schedule B to the Harmonized System (HS) commodity codes. Previously, schedule B was matched to the Tariff Schedule of the United

States.[7] The major revision did affect classification at the mid-level of aggregation reported in the OM series. Since the numerical value of the schedule B codes changed, concordance before and after 1989 must be done using written descriptions. The extent to which changes in data in 1988 and 1989 reflect changes in schedule B codes is not known, though the Census does not seem particularly concerned about it.

In 1990 a second major change occurred. Previously Canadian data indicate substantially larger imports from the United States than the United State's export data. That there is a difference is not surprising as imports are under more scrutiny than exports. It is the size of the error that is surprising. The difference in 1989 is close to $15.6 billion (*Merchandise Trade Statistics* 1994) or nearly 20% of documented U.S. exports. To correct this problem, in 1990 the Census began using the import statistics collected by Canada. There is a huge increase in documented exports to Canada when using the Canadian import data rather than the United States export data reflecting the greater scrutiny of import inspectors. Because Canada is such a large export destination, switching to import data results in large swings in *state export totals*, not just state Canadian exports. The Census warns of comparability issues pre- and post-1990. To use the entire OM time series, one must decide how to assign the discrepancy between imports and exports for the years 1987–1989. Because states do not export to Canada evenly, and most exports to Canada originate in bordering states, there is potential for large errors if values are not assigned properly. Cronovich and Gazel (1999) modify the data by assigning the difference in reported U.S. exports and Canadian imports according to state and subsector export shares in 1990 data. They find substantially altered data of pre-1990 years with this assignment procedure.

The final problem with the time series is the change from SIC to NAICS. In the early years of the OM series, from 1987–2000, exports are listed by the 1987 SIC revision. Data from the years 1997–2002 are available by the 1997 NAICS. Data is listed by the 2002 NAICS revision from 2002 onwards. Differences between NAICS97 and NAICS02 is minimal at the three digit level, so this break in the time series is not serious. The break after 2000, when SIC data is no longer reported, is serious because the differences between SIC87 and NAICS97 are substantial. Table 13 shows a

[7] The Tariff Schedule of the United States, administered by the U.S. International Trade Commission, was adjusted to match HS in 1989 also.

rough concordance for the level of aggregation reported by the OM series. Though not perfect, the rough concordance in table 13 gives an idea of the match between the two classification systems. One may ease into the switch in classifications because both SIC87 and NAICS97 are reported for the years 1997–2000. Therefore direct comparisons may be made. Export shares averaged over several years may be calculated to assign data in the older SIC categories into newer NAICS codes. The change from SIC to NAICS has no effect on total exports to each destination, so the problem is moot if one wished to use state and destination information.

4 Quality of OM Data as Representative of Origin of Production

Coughlin and Wall (2003) refer to the OM data as the best available source of state export data. Of course, it is the only source to provide both state origin and destination. Nonetheless, a data source of poor quality should not be used even if it is the only data source. In this section, diagnostic tests asses the quality of the OM data.

To avoid confusion, precise meanings are given to two words previously used loosely.

Definitions. *The **representativeness** of data set A to another, B, is the degree of agreement of results for a given class of models, when A replaces B. Synonymous with **substitutability**.*

*The **quality** of a data set is its representativeness of the "true" data set.*

Concerns about the quality of the OM data arise only over the assignment and interpretation of *state* export values, not national values. That is, everyone agrees the OM data on total United States exports, U.S. exports by subsector, or U.S. exports to a particular destination is the most accurate available.[8] This data is used extensively in practice. The issue is whether the OM data represents the OP of *state* exports.

The following steps determine the quality of the OM data. Section 4.1 considers the Census's warning about the OM data and finds agricultural and mining products data in the OM do not represent OP. Next section 4.2 introduces another Census state export data set, the AR-1 series on manufacturing. Though this set does not contain destinations, the Census recommends it for

[8]Import data of some foreign destinations is considered more accurate than U.S. export data.

OP of manufactured exports. Despite flaws causing the AR-1 to differ from the "true" OP, the AR-1 will be the object of comparison to test for the quality of the OM. Hence the OM data may have high quality despite not representing AR-1. Likewise, the OM data may represent AR-1 but be low quality. Since neither scenario may be tested, in practice the OM is considered high quality if it is largely substitutable for AR-1.

Sections 4.3 and 4.4 test the representativeness of OM to AR-1 data, the former for *overall* quality and the later for quality of *individual* subsectors and states. Each section applies two tests: matched pairs and limits of agreement for overall quality and fixed effects regression and descriptive tests for individual quality. The discussion on quality ends with section 4.5 which recalls two studies comparing the OM and AR-1 sets using older data. Their conclusions are similar though their methods are not.

4.1 Census Warnings and the Poor Quality of Agriculture and Mining

The determination of OM quality begins with the Census.[9] The Census warns, "The [OM] series DOES NOT represent the production origin of U.S. export merchandise."[10] The Census, however, also acknowledges, "There are nonetheless many cases when the state origin of movement and the state of production happen to be the same"(*Guide to State Export Data*; GSED).

The Census explicitly warns about agricultural and mining exports. "Most affected [by consolidation at ports] is the allocation of exports of farm products, minerals, and other bulk commodities—virtually all of which are sold abroad by intermediaries" (GSED). This warning is due to observations in which OM reports large exports in subsectors for states known to have little production in those subsectors. For example, in 2003 Louisiana is credited with exporting twelve times more value of agricultural and animal products (NAICS 111 plus 112; $9,388 million) than its Gross State Product ($790 million) in the combined subsector. Other examples in agriculture and mining are not hard to find.

[9]The International Trade Administration (ITA) administers the "Exporter Database" which uses OM data, and is a good source of information about it. References to "Census" include the ITA though it is a separate agency in the Department of Commerce.

[10]`http://www.census.gov/foreign-trade/aip/elom.html`. Emphasis in original.

The Census is more optimistic about manufacturing. "The impact on manufactured exports is much more limited, due to the fact that intermediaries account for only about one-third of U.S. exports of manufactures" (GSED). Furthermore there are no red flags for manufacturing as there is for agriculture. Nonetheless, for data on OP of manufactured exports, the Census recommends another state export data set. This other data, the AR-1 series, is the best data available for OP.

The Census is silent on the other sectors such as prepackaged software. These remaining subsectors are extremely small, and so are discarded without much loss.

4.2 AR-1 Data on State Exports

"Exports from Manufacturing Establishments" (AR-1) is a Census data set compiled using surveys of manufacturing firms.[11] The AR-1 series has data on manufactures only, and for the fifty U.S. states and the District of Columbia; no comparison on agriculture, mining, or other subsectors, or on Puerto Rico or the U.S. Virgin Islands is possible.

The AR-1 series published annually from 1983 to 1991. It subsequently published for the years 1997 and 2000–2002 with the possibility of more recent years forthcoming. There is overlap between the AR-1 and OM series in the years 1987–1991, 1997, and 2000–2002. The AR-1 series uses data from the Annual Survey of Manufactures (ASM) or the Economic Census to determine the physical location of exporters. The ASM asks manufacturers to "Report the value of products shipped for export. Include direct exports and products shipped to exporters or other wholesalers for export... Do not include products shipped for further manufacture, assembly, or fabrication in the United States." Hence AR-1 data are free-on-board (f.o.b.) plant values.

Though asked for indirect exports, the export data from the ASM is frequently below export totals gathered by the SED. In addition to problems of non-response and under-coverage (roughly 25% of plants receive an ASM survey), many manufacturers do not know the final destination of their products due to the actions of independent wholesale or retail traders. Thus totals do not match those collected by the SED.

[11] Another data set, "Selected Characteristics of Manufacturing and Wholesale Establishments that Export" (AR-2) is not used for determining if the OM data represents OP as it overlaps with the OM data for only two years, 1987 and 1992. Details of AR-2 are available in appendix A.

Table 1. Data comparison: OM vs. AR-1

Data	Years	Exporter locations	Destinations	Source	Measured	Censored
OM	1987–present	50 U.S. States, D.C., Puerto Rico, U.S. Virgin Islands	242	Customs	F.a.s.	No
AR-1	1983–1991, 1997, 2000–2002	50 U.S. States, D.C.	0	Survey	F.o.b.	Yes

Note: AR-1 contains data only on twenty-one manufacturing subsectors (NAICS 31x and 32x) whereas OM contains data on manufacturing as well as agriculture, mining, and various other subsectors.

The Census uses the *U.S. International Trade in Goods and Services* (unpublished, though some data disseminated as the FT-900 series) to make adjustments to the ASM data to better match U.S. export totals as collected by the SED. The Census takes the discrepancy between reported exports in ASM and measured exports from the SED for each state and subsector and allocates it to form the published AR-1 data.[12] The allocation scheme assumes a direct relationship between the discrepancy and the geographic distribution of the total value of shipments and reported exports. Furthermore, before allocation, the Census converts the discrepancies from f.a.s. into f.o.b. values, using estimates from the *Input-Output Accounts* of the Bureau of Economic Analysis, published every five years. These conversions apply with regard to subsector, but not state. Details of the allocation scheme are given in *Exports From Manufacturing Establishments*. The allocation scheme requires many estimates and conversions where errors occur. The use of the SED in the estimation of the AR-1 does not prevent its use for comparison to the OM because the SED export data is the best available for national export totals.

Unlike the OM, AR-1 is censored by the Census. AR-1 data on many subsectors are not released either due to disclosure concerns of publishing identifying information or because the data did not meet quality standards. Furthermore the Census does not report exports of less than $100,000. Nondisclosure and truncation apply more frequently to small states and subsectors than large, but are not problems for subsector and state totals. Note, however, that by construction the U.S. total from the AR-1 data is no less than the f.o.b. valued total collected from the SED.

Though the AR-1 data is considered the best data available for OP, it has flaws. As reported in Bernard and Jensen (1995), the ASM under counts exports from small subsectors: printing and

[12]The Census only allocates the discrepancy when the values from the SED are greater than those in the ASM. In the unusual case where ASM values are greater, the data is not modified.

publishing, leather, apparel, and wood products. Other drawbacks of AR-1: the data set is survey based so it is subject to sampling and non-response errors from the "true" state of production values, errors in the estimates, only manufacturing subsectors are available, the destination of exports is not available, and the data is subject to disclosure concerns. The under counting of small subsectors, sampling and non-response issues, as well as the differences in f.a.s. to f.o.b. values should be kept in mind when viewing the comparison of the AR-1 and OM sets.

4.3 Overall Quality of the OM Data

Two diagnostic tests determine the substitutability of the OM and AR-1 data. These are matched pairs and limits of agreement. Matched pairs compares percent or mean differences in the data sets across different dimensions of data. Though the results for 1991 and 1997 are given; the focus is on 2001 since no WISER adjustments are made to it. The quality of the OM data should be improved over earlier years, 1991 and 1987, because of the reduction of reporting errors and the higher quality of Canadian exports.

Values reported in AR-1 are not, and should not be, equal to those in the OM. One reason is the AR-1 has nondisclosure issues. The OM does not. Another reason is the OM data contain f.a.s. values whereas the AR-1 data are f.o.b. In 2001 the difference in national levels is $70 billion due to f.a.s vs. f.o.b. Furthermore, if interior states have higher transportation costs than port states, then one expects the percent difference between OM and AR-1 data to be greatest for interior states. One may think converting the OM data to f.o.b. would increase the ease of comparison. This is true, but it also increases measurement error due to the three digit level of aggregation and the fact transportation costs by state are not known. Thus despite using different units of measurement, the data will be compared as is.

The method of matched pairs uses the following hypothesis:

$$\log OM_{x,t} = \log trans_{x,t} + \log AR_{x,t} + \varepsilon_{x,t} + \delta_{s,t} \tag{1}$$

where $trans$ is the inland freight transportation cost to convert f.o.b. into f.a.s., ε is the logged random error and δ is the logged error due to consolidation or geographic locations. The subscript x

refers to either the subsector or state dimension of the data. This model emphasizes transportation costs; both random and consolidation errors are multiplicative.

A note is in order about the types of tests *not* chosen. It is not appropriate to compare the data sets by testing if the underlying distributions are the same since the size and rank of the states matter. This includes a paired t-test for matching means. Also, coefficients of correlation are not meaningful as they reflect the strength of the linear relation between the data sets, not whether one is substitutable for the other.[13] For example, both an interior state and a port state could have high correlations, though the linear associations would have slope less than one for the interior state and slope greater than one for the port state. The differences in slope reflect the problems of consolidation despite high correlations in both states. Furthermore, correlation is sensitive to outliers and thus favors matching the largest NAICS subsectors in each state at the expense of many smaller subsectors. See Altman and Bland (1983) for a full discussion of the problem of agreement.

Finally, the AR-1 is considered the best data on OP, but is known not to be the "true" OP. Neither the OM nor the AR-1 are samples from a "true" population, but rather the population measured imperfectly. Hence quality tests look for comparability or substitutability between the OM and AR-1. This means substitutability of OM and AR-1 may not imply OM is OP and the non-substitutability may not imply OM is not OP.

Matched Pairs

The method of matched pairs finds the difference between the same observation (either by state, NAICS, or state-NAICS) from the OM and AR-1. This difference is then divided either by AR-1 observations or by the average of OM and AR-1 (henceforth referred to as "mean difference"). Mean difference has the advantage of handling cases when AR-1 is zero. However since the data sets differ by transportation costs as well as the two types of error, the mean is not reliable as the "true" value as measured by either f.a.s. or f.o.b. units. Thus calculations are primarily performed using percent difference.

[13] For those who are nonetheless curious: correlations by state, NAICS, and state-NAICS are all greater than 0.90 for 2001 and 1997.

It is clear from section 4.2 the difference in OM and AR-1 levels should not be zero when comparing either by state or subsector. This is due, in part, to the differences in export valuation: f.a.s. for OM and f.o.b. for AR-1. Thus the levels of differences are not useful for overall quality. Rather quality is measured by the *variation* among percent differences of the matched pairs. However simply finding the variation of differences across states is uninformative since there needs to be a criterion to determine if the observed variance is large enough to claim the OM data is not appropriate to use as OP. Thus the variance of percent differences cut by subsectors is reported first as a guideline for comparison with the state slice.

Using (1), matched pairs—with percent differences given by $(OM - AR1)/AR1 * 100$—compares the standard deviation (std. dev.) in $trans_{state} * e^{\varepsilon+\delta} - 1$ to the std. dev. in $trans_{NAICS} * e^{\varepsilon} - 1$. The reason e^{δ} does not appear on the subsector detail is because, by construction, AR-1 matches OM up to the transportation conversion. Sample std. dev. measure the variance rather than the population std. dev. because though the data are the population, it is the population measured with error.

Table 2 lists percent and mean differences between the OM and AR-1 data sets on a three digit NAICS basis for two years: 2001 and 1997. 1991 is available in the online appendix.

The "OM" columns in table 2 are the sum of sales of the fifty U.S. states. Exports from the District of Columbia, Puerto Rico, and the U.S. Virgin Islands or of unknown origin are not included. Thus the same data are used on both this subsector detail and the forthcoming state detail, however, the OM value for the U.S. differs from other sources of total U.S. manufacturing exports. The "AR1" columns are the values reported in the subsector detail of *Exports From Manufacturing Establishments*. The sum of states by subsector in the AR-1 does not equal this value due to nondisclosure. The cases of negative differences are due to the removal of exports with unknown state of origin of movement. Column "UK" is the percent of a subsector's total exports that are of unknown origin of movement. If these unknown exports are added to the subsector totals, all differences are positive.

For 2001 the std. dev. of percent differences is 10.94. For 1997 the std. dev. is substantially larger at 43.89 due to printing and publishing (NAICS 323), a relatively small subsector. If printing and publishing is removed, std. dev. becomes 10.89 in 2001 and 11.37 in 1997. Keep these std. dev.

Table 2. Comparison of OM to AR-1 by NAICS

Description	NAICS	2001					1997				
		OM	AR-1	UK	Percent difference	Mean difference	OM	AR-1	UK	Percent difference	Mean difference
		(Billions $2001)		(%)	(hundreths)		(Billions $1997)		(%)	(hundreths)	
Food prdcts.	311	25.59	23.75	3.12	7.77	7.48	22.31	23.34	15.24	−4.43	−4.53
Bev. & tbc.	312	4.30	3.86	1.26	11.38	10.77	6.38	6.05	4.06	5.40	5.26
Txt. & fbrcs.	313	7.05	6.50	3.90	8.34	8.01	4.82	5.10	14.47	−5.37	−5.52
Txt. mill prdcts.	314	1.94	1.85	7.08	4.57	4.46	1.89	1.97	15.80	−3.81	−3.88
Apparel	315	6.57	5.61	3.57	17.15	15.80	7.84	7.36	7.79	6.59	6.38
Leather	316	2.57	2.00	4.82	25.19	11.97	2.19	2.13	17.40	2.67	2.64
Wood prdcts.	321	3.96	3.19	3.38	21.46	15.96	4.43	4.40	24.55	0.76	0.76
Paper	322	14.01	12.30	3.09	13.92	13.02	12.59	12.61	15.83	−0.18	−0.18
Publishing	323	4.81	4.61	6.06	4.42	4.32	4.41	1.47	10.19	200.17	100.04
Petro. & coal	324	7.87	7.72	4.24	1.84	1.82	6.43	6.67	8.41	−3.48	−3.54
Chemicals	325	70.96	61.74	2.04	14.93	13.89	62.52	57.37	6.55	8.97	8.58
Plscts. & rub.	326	15.76	13.04	4.46	20.89	18.92	12.63	11.43	10.90	10.48	9.96
Nonmetal mnrl.	327	7.44	6.31	3.80	18.13	16.62	5.77	5.36	11.12	7.69	7.41
Primary metal	331	18.44	16.24	4.39	12.57	12.71	19.07	18.07	9.19	5.53	5.38
Fab. metal	332	18.99	17.02	7.81	11.60	10.96	15.97	16.36	20.02	−2.36	−2.38
Machinery	333	78.14	62.78	4.10	24.46	21.79	19.07	18.07	10.79	5.53	13.20
Comps. & elect.	334	157.67	116.93	3.60	34.85	29.68	144.22	114.97	4.96	25.45	22.57
Elect. eqpmnt.	335	23.67	19.14	4.61	23.69	21.18	20.77	18.55	9.32	11.98	11.30
Trans. eqpmnt.	336	126.41	117.75	2.53	7.35	7.09	109.34	108.76	7.81	0.53	0.53
Furniture	337	2.25	2.24	14.73	0.73	0.72	2.03	2.23	22.84	−9.06	−9.49
Misc. mnfcts.	339	29.94	16.81	5.87	42.38	34.97	18.47	13.24	7.96	39.47	32.96
All mnfcts.	31–32	622.36	521.39	3.61	19.37	17.66	558.79	502.88	9.62	11.12	10.53
Std. dev.		–	–	–	10.94	10.93	–	–	–	43.89	22.98

Sources: OM data from WISER; AR-1 data from *Exports From Manufacturing Establishments*, Census.

Notes: UK is the value of exports by subsector not assigned to any state as a percent of state sum of exports; Percent difference is $(OM - AR1)/AR1$; Mean difference is $\frac{OM-AR1}{(OM+AR1)/2}$; Std. dev. is $\sqrt{\frac{\sum(x-\bar{x})}{N-1}}$ where x is the percent or mean difference of the matched pair, $\bar{x}$ is the average of the percent or mean differences and N is the number of observations.

in mind when looking at cuts of the data by state, in table 3 below.[14]

Table 3 compares total manufacturing exports reported in the OM and in the AR-1 by state for four years: 2001, 1997, 1991, and 1987. The left half of table 3 shows percent differences of total manufacturing export sales between the OM and AR-1 data. Again, the OM value is the sum of subsectors for each state without assignment of quantities of unknown origin whereas the AR-1 data is the reported state total, not the sum of subsectors in each state.

In 2001 the std. dev. is 51.21. It is 197.51 in 1997. Wyoming is the cause for the large std. dev. in 1997. Without Wyoming std. dev. drops to 52.44 in 1997 and 46.08 in 2001. Std. dev. in 1991 is 681.63 but is 60.68 without Wyoming. In 1987 it is 78.29 and 46.53 when Wyoming is removed.

[14]If the data from the OM's "unknown" state are included, the percent differences all increase by various amounts according to the size of the subsector's unknown share. The standard deviations increase slightly to 11.14 and 47.48 in 2001 and 1997 respectively.

Table 3. Comparison of OM to AR-1 by state

State	Percent difference in OM and AR-1 levels				Percent difference in OM and AR-1 shares			
	2001	1997	1991	1987	2001	1997	1991	1987
AK	12.13	34.97	−25.76	70.41	−6.06	22.10	−29.18	52.81
AL	−9.22	−25.42	−28.28	−11.31	−23.95	−32.53	−31.58	−20.47
AR	−45.95	−48.98	−53.26	−52.99	−54.72	−53.85	−55.41	−57.84
AZ	47.78	92.64	49.21	32.92	23.80	74.27	42.33	19.19
CA	39.98	34.56	28.24	32.41	17.27	21.73	22.33	18.73
CO	20.40	−11.88	−0.17	−10.72	0.87	−20.29	−4.77	−19.97
CT	7.61	−19.38	−31.86	−34.68	−9.85	−27.07	−35.00	−41.43
DE	134.98	117.35	43.08	62.01	96.86	96.62	36.49	45.28
FL	177.06	118.78	71.45	99.93	132.11	97.92	63.55	79.29
GA	−6.42	−8.00	−5.54	−5.08	−21.60	−16.77	−9.90	−14.88
HI	73.97	−90.90	−67.18	−12.44	45.74	−91.77	−68.69	−21.48
IA	−39.76	−38.58	−40.86	−31.20	−49.53	−44.44	−43.59	−38.30
ID	−36.88	−66.22	−55.19	−39.67	−47.12	−69.44	−57.23	−45.90
IL	26.91	7.28	−4.22	−2.49	6.32	−2.95	−8.63	−12.56
IN	−20.87	−28.83	−35.97	−17.96	−33.71	−35.62	−38.92	−26.43
KS	−14.04	−3.71	−34.80	−22.06	−27.99	−12.89	−37.81	−30.11
KY	−1.19	−26.46	−43.97	−33.56	−17.22	−33.48	−46.55	−40.42
LA	11.86	23.23	30.31	72.07	−6.29	11.48	24.30	54.30
MA	18.64	18.29	4.38	27.51	−0.61	7.01	−0.43	14.34
MD	15.12	15.44	14.40	−2.43	−3.56	4.43	9.12	−12.50
ME	−32.64	−21.11	−34.69	−18.42	−43.57	−28.63	−37.70	−26.84
MI	14.57	24.67	24.00	41.94	−4.02	12.78	18.29	27.29
MN	17.99	15.17	8.15	−18.66	−1.15	4.19	3.16	−27.06
MO	−34.72	−47.01	−40.55	−44.63	−45.31	−52.06	−43.29	−50.34
MS	8.09	−53.93	−40.78	−34.45	−9.45	−58.32	−43.51	−41.22
MT	−25.47	−23.98	45.52	23.07	−37.56	−31.23	38.82	10.36
NC	−10.69	−15.02	−28.95	−13.62	−25.18	−23.12	−32.22	−22.54
ND	5.69	17.96	−12.81	−2.34	−11.46	6.71	−16.83	−12.42
NE	−30.10	−39.88	−43.24	−9.78	−41.44	−45.61	−45.85	−19.10
NH	−11.36	−54.37	−27.93	−28.04	−25.74	−58.72	−31.25	−35.47
NJ	65.22	56.75	35.55	59.39	38.42	41.81	29.30	42.93
NM	−48.26	−72.12	−54.29	−27.72	−56.66	−74.78	−56.39	−35.18
NV	32.85	40.13	104.09	109.51	11.30	26.76	94.69	87.88
NY	72.27	50.86	46.07	48.97	44.32	36.47	39.34	33.59
OH	−11.89	−19.13	−29.37	−31.06	−26.18	−26.84	−32.63	−38.18
OK	−42.20	−34.66	−12.05	−26.86	−51.58	−40.89	−16.10	−34.42
OR	−7.55	−2.14	−18.22	8.18	−22.55	−11.47	−21.99	−2.99
PA	−6.69	−8.19	−14.97	−14.63	−21.83	−16.94	−18.89	−23.45
RI	−19.18	−39.83	−38.87	−40.96	−32.29	−45.57	−41.68	−47.06
SC	−18.59	−21.45	−31.37	−33.24	−31.80	−28.94	−34.54	−40.13
SD	−60.56	−56.02	−37.11	−79.28	−66.96	−60.21	−40.01	−81.42
TN	−2.70	−15.39	−32.29	−35.26	−18.49	−23.46	−35.41	−41.95
TX	112.38	82.75	67.18	61.34	77.92	65.32	59.47	44.68
UT	0.27	−7.08	23.86	28.50	−15.99	−15.94	18.15	15.23
VA	−2.20	9.49	23.73	29.92	−18.07	−0.95	18.03	16.51
VT	−11.64	181.77	322.82	17.66	−25.98	154.90	303.34	5.51
WA	−7.12	7.02	−2.74	8.78	−22.19	−3.19	−7.23	−2.45
WI	−17.09	−19.47	−26.86	−14.80	−30.54	−27.15	−30.23	−23.60
WV	−14.09	−20.53	−35.25	−18.31	−28.03	−28.11	−38.24	−26.75
WY	171.40	1347.69	390.11	475.19	127.37	1209.63	367.52	415.79
US	19.37	10.55	4.83	11.52	19.37	10.55	4.83	11.52
Std. dev.	51.21	197.51	81.63	78.29	42.90	177.75	77.87	70.21

Sources: Author's calculations for 2001 and 1997 use OM data from WISER and AR-1 data from *Exports from Manufacturing Establishments*, Census; Author's calculations for 1991 use OM data from Feenstra (1997) and *Exports from Manufacturing Establishments*, Census; Data for 1987 from Coughlin and Mandelbaum (1991).

Notes: Percent difference is $(OM - AR1)/AR1 * 100$; Std. dev. is $\sqrt{\frac{\sum(x-\bar{x})}{N-1}}$ where x is the percent difference of the matched pair, $\bar{x}$ is the average of the percent differences and N is the number of observations.

Table 4. Standard deviations of matched pairs

Year	Percent difference				Mean difference			
	NAICS		State		NAICS		State	
	All	No 323	All	No WY	All	No 323	All	No WY
2001	10.94	10.89	51.21	46.08	9.13	9.07	38.81	36.87
1997	43.89	11.37	197.51	52.44	22.98	11.37	56.25	50.23
1991	17.92	18.38	81.63	60.68	16.18	16.58	47.74	43.56

The right half of table 3 shows the percent differences between export shares. For the OM share, the total value of exports from each state reported in the OM series is divided by the U.S. total—the sum of the fifty states. Shares sum to one. The share percent differences are systematically lower than the level percent differences, though the signs are roughly the same. Therefore the right half of table 3 supports any conclusions drawn from the left half.

The std. dev. of mean differences are similar to the percent differences. For export levels, they are 38.81, 56.25, 47.73, and 44.38 in 2001, 1997, 1991, and 1987.

Table 4 summarizes the findings in tables 2 and 3. Notice, the std. dev. of percent differences for the state detail are greater than for the subsector detail by roughly five times for all years. Though five times greater than the subsector detail, the state std. dev. are not alarmingly large. Because of inland transportation costs and measurement error in AR-1, even if consolidation is not a problem, the std. dev. on the state slice should be larger than the subsector slice. Thus consolidation, e^{δ}, at most causes five times more std. dev. but is likely far less. The OM data, as an entire set, is in sufficient agreement with the AR-1 data to be used as OP of exports.

Limits of Agreement

If the data from OM is within some acceptable limits around AR-1 ninety-five or ninety percent of the time, then the sets are substitutable. Table 5 shows the 95% and 90% limits of agreement for the OM and AR-1 data cut by subsector, state, and subsector-state pairs and the data needed to construct them. The log transformation from (1) causes the limits of agreement in table 5 to represent the coefficient applied to AR-1 to make it OM.

At first glance, the limits in table 5 look wide suggesting the two data sets are not substitutable. Keep in mind several points. First, OM and AR-1 are measured in different units. Second, the

Table 5. Limits of agreement

Year	Slice	N	Mean	Std. dev.	U95	L95	U90	L90
	NAICS	21	0.144	0.092	1.40	0.95	1.35	0.99
2001	States	50	0.009	0.406	2.28	0.45	1.99	0.51
	St-NCS[a]	824	0.087	0.878	6.10	0.20	4.63	0.26
	NAICS[a]	21	0.099	0.250	1.86	0.66	1.70	0.72
1997	State[a]	50	−0.082	0.694	3.73	0.23	2.95	0.29
	St-NCS[a]	965	0.103	1.106	9.68	0.13	6.83	0.18
	SIC	20	0.031	0.163	1.45	0.73	1.37	0.78
1991	State[a]	50	−0.094	0.518	2.57	0.32	2.16	0.38
	St-SIC[a]	631	−0.073	0.846	4.88	0.18	3.74	0.23

Sources: OM data from WISER; AR-1 data from *Exports From Manufacturing Establishments*, Census.

Note: Reported confidence intervals are antilogs of the true intervals.

[a]Null hypothesis of normally distributed differences rejected at 5% level using a Shapiro-Wilks test indicating limits of agreements biased.

state limits are less than twice the subsector cuts, indicating consolidation cannot be too large of a problem. Third, the log transformation loses information because there are many zero observations in the AR-1. Including nondisclosed observations, there are 193, 27, and 219 zeros in 2001, 1997, and 1991 respectively. Fourth, some slices of data reject the hypothesis of normally distributed errors despite the log transformation. Since values from the normal are used in constructing the limits of agreement, the true limits may differ from those reported.

The matched pairs and limits of agreement show there is more variation, five times greater std. dev., in the state slice of data than the subsector slice. Some of this additional variance is due to inland freight transportation costs and measurement error in AR-1. The remaining variance is due to consolidation. Since the std. dev. between cuts is well below an order of magnitude, the variance due to consolidation is not large enough to reject the OM data as not representative of AR-1. Though consolidation is present, the *overall* OM data may be used for AR-1 and OP.

4.4 Quality of Individual Subsectors and States in the OM data

Though the *overall* OM data represents OP, there may be some individual subsectors or states that do not. Regression is the primary tool to identify the subsectors or states that are not substitutable in the two data sets. Regression, however, is not the end-all because the f.o.b. to f.a.s. conversion confounds with the consolidation error. Furthermore the regression has biased parameter estimates

since AR-1 has measurement error. Therefore a simple descriptive test further highlights subsectors or states of low quality.

Scatterplots provide a visual test of the substitutability of OM and AR-1 data. They plot the un-transformed OM data against the AR-1 data indicating the linear relationship between the data. Included with the scatterplots are a regression line and a line with slope of 1.15. The line of slope 1.15 serves as a rough guide for transportation cost mark-up. Appendix B contains scatterplots for a number of slices of the data, including individual subsectors and states for 2001. The scatterplots give some idea of how the two sets compare visually, but are not used in the analysis of quality.

Fixed Effects Regression

Regression is the standard method for identifying subsectors or states of low quality. The model:

$$\log OM_{x,t} = \sum_{x} \beta_{0_{x,t}} D_x + \beta_{1_{x,t}} \log AR1_{x,t} + \varepsilon_{x,t} \tag{2}$$

where t is the time index, and x is either the subsector or the state index depending if the group is state or subsector. D_x is a dummy taking the value of one for state or subsector x and zero otherwise. Using logs causes the transportation costs and consolidation and random error to be multiplicative as well as normalizing errors and reducing heteroskedasticity. Even so, heteroskedasticity remains a problem so the regression uses robust standard errors.

This regression uses fixed effects to correct for an unobserved but correlated variable such as state transportation cost. If fixed effects are not used, the least squares estimator is biased and inconsistent. Because transportation costs change over time, the fixed effects are not constant. Hence the regression does not pool years but uses one year at a time. Given data is not pooled by year, no subsector-state interaction dummies are used because there are not enough observations.

Equation (2) corresponds to (1); the constant term $\beta_{0_x} = \log trans_x + \delta_x$. Separately identifying these is not attempted. Rather, constants far from zero suggest, but do not confirm, consolidation problems. Slope coefficients near one suggest OM is substitutable for AR-1. Three warnings: since AR-1 is a population measured with error, β_1 may not be close to one due to attenuation rather than non-agreement of data sets, censorship and truncation of the AR-1 data result in the removal

of some observations, and the regression hides states underestimated by OM. This is because for these states, $\delta < 0$ which offsets $\log trans$ for a net value near zero.

Tables 6 and 7 are the fixed effects results for 2001 and 1997. Since the subsector data in each set matches up to the transportation cost conversion, $\beta_{0_{NAICS}} = \log trans_{NAICS} + 0$ thus establishing guidelines for comparison to the state regression.

In table 6, β_1 slope coefficients are acceptably close to one given the attenuation from measurement error. The claim is confirmed by switching the OM and AR-1 data in the regression. Intercept estimates are almost all significant. The exceptions are apparel (NAICS 315), leather (316), printing and publishing (323), and petroleum and coal products (324) in 2001 and beverages and tobacco (312) in 1997. Suspect subsectors either have large intercepts or intercepts that change from one year to the next. These are beverages and tobacco (312), textiles and fabric (313), printing and publishing (323), chemicals (325), machinery (333), computers (334), and transportation equipment (336).

A slight variation of the fixed effects regression above gives intercept estimates as deviations from the average (Suits 1984). This exercise (not reported) largely reflects the findings above, but adds leather (316) to the list of problematic subsectors.

In table 7 several states have large intercepts or intercepts changing greatly from 1997 to 2001. These are: Alaska, Arizona, California, Delaware, Florida, New Jersey, New York, Texas, and Wyoming. Large state intercepts indicate states where OM is greater than AR-1, but β_0 hides states where the opposite occurs.

Because of the attenuation from measurement error in AR-1 and the fact it hides states underestimated by OM, fixed effects regression cannot identify all individuals that are poor quality. Next a descriptive test finds low quality individuals using percent differences.

Descriptive Test

Though it is the variation of differences between OM and AR-1 needed to test for overall quality, the levels of differences inform about individual quality. Return to table 2. In 2001 there are no subsectors with a percent difference greater than 75% or less than -45%. These bounds are roughly four times 19.37%, the percent difference in all manufactures. Though somewhat arbitrary,

Table 6. Subsector fixed effects

	Stat.	2001		1997	
	N	824		965	
	R^2	0.98		0.97	
	β_1	0.84+	(0.02)	0.80+	(0.02)
Description	NAICS	β_0	s.e.	β_0	s.e.
Food prdcts.	311	0.91∗	0.15	0.90∗	0.16
Bev. & tbc.	312	1.22∗	0.26	0.62	0.34
Txt. & fbrcs.	313	0.89∗	0.25	1.37∗	0.22
Txt. mill prdcts.	314	0.44∗	0.17	0.52∗	0.21
Apparel	315	0.25	0.22	0.47∗	0.19
Leather	316	0.52	0.26	0.63∗	0.21
Wood prdcts.	321	1.00∗	0.14	0.66∗	0.13
Paper	322	0.91∗	0.15	0.83∗	0.15
Publishing	323	0.22	0.17	1.30∗	0.21
Petro. & coal	324	0.43	0.28	0.76∗	0.27
Chemicals	325	1.29∗	0.16	1.50∗	0.19
Plscts. & rub.	326	0.99∗	0.15	1.08∗	0.15
Nonmetal mnrl.	327	0.92∗	0.15	0.79∗	0.14
Primary metal	331	0.75∗	0.17	0.98∗	0.20
Fab. metal	332	0.73∗	0.17	0.82∗	0.15
Machinery	333	1.14∗	0.16	1.24∗	0.17
Comps. & elect.	334	1.21∗	0.21	1.35∗	0.18
Elect. eqpmnt.	335	0.74∗	0.18	0.95∗	0.17
Trans. eqpmnt.	336	1.17∗	0.18	1.36∗	0.21
Furniture	337	0.70∗	0.13	0.47∗	0.18
Misc. mnfcts.	339	0.82∗	0.15	0.93∗	0.15

Note: F-values are 5.01 and 4.01 respectively, rejecting the hypothesis that β_0 is equal across groups.

Sources: Author's calculations using OM data from WISER and AR-1 data from *Exports From Manufacturing Establishments*, Census.

Notes: Fixed effects regression:

$$\log OM_{x,t} = \sum_x \beta_{0_{x,t}} D_x + \beta_{1_{x,t}} \log AR_{x,t} + \varepsilon_{x,t}$$

where D_x is a dummy variable. Standard errors are robust. Non-disclosed and truncated observations are not included.

(+) indicates significantly different from one at 5% level; (*) indicates significance at the 5% level.

Table 7. State fixed effects

Stat.	2001		1997	
N	824		965	
R^2	0.98		0.97	
β_1	0.83+	(0.02)	0.76+	(0.02)
State	β_0	s.e.	β_0	s.e.
AK	0.41	0.10	1.78∗	0.50
AL	0.65∗	0.18	1.17∗	0.17
AR	0.17	0.23	0.70∗	0.18
AZ	1.55∗	0.25	1.83∗	0.23
CA	1.75∗	0.17	2.04∗	0.17
CO	0.57∗	0.19	0.83∗	0.23
CT	0.94∗	0.22	1.32∗	0.19
DE	1.50∗	0.27	1.24∗	0.22
FL	1.94∗	0.16	2.18∗	0.16
GA	0.99∗	0.16	1.44∗	0.17
HI	−0.30	0.60	0.47	0.42
IA	0.65∗	0.17	0.80∗	0.22
ID	0.66∗	0.17	0.76∗	0.32
IL	1.18∗	0.19	1.68∗	0.18
IN	0.76∗	0.25	1.06∗	0.24
KS	0.71∗	0.15	1.03∗	0.16
KY	0.80∗	0.19	0.94∗	0.20
LA	0.97∗	0.17	1.72∗	0.36
MA	1.02∗	0.19	1.53∗	0.17
MD	1.06∗	0.31	1.41∗	0.24
ME	0.46∗	0.23	0.64∗	0.19
MI	1.00∗	0.16	1.44∗	0.17
MN	0.94∗	0.21	1.37∗	0.16
MO	0.71∗	0.17	0.90∗	0.18
MS	0.73∗	0.23	0.74∗	0.20
MT	0.68∗	0.27	0.35	0.29
NC	0.98∗	0.15	1.32∗	0.15
ND	0.95	0.54	0.36	0.41
NE	0.25	0.25	0.55∗	0.23
NH	0.51∗	0.19	0.70∗	0.15
NJ	1.39∗	0.18	1.61∗	0.17
NM	0.39	0.32	−0.12	0.25
NV	0.72∗	0.28	0.83∗	0.21
NY	1.70∗	0.19	1.93∗	0.19
OH	0.95∗	0.16	1.37∗	0.15
OK	0.33	0.19	0.53	0.27
OR	0.94∗	0.21	1.36∗	0.19
PA	0.92∗	0.14	1.23∗	0.19
RI	0.51∗	0.19	0.38	0.17
SC	0.66∗	0.20	0.90∗	0.28
SD	−0.16	0.21	0.54	0.30
TN	0.76∗	0.17	1.16∗	0.19
TX	2.15∗	0.19	2.41∗	0.19
UT	0.59∗	0.26	0.62∗	0.32
VA	0.85∗	0.16	1.35∗	0.18
VT	0.70∗	0.21	1.18∗	0.26
WA	1.05∗	0.14	1.32∗	0.17
WI	0.86∗	0.15	1.10∗	0.14
WV	−0.01	0.35	−0.38	0.39
WY	1.47∗	0.28	0.36	0.66

Note: F-values are 6.06 and 6.23 respectively, rejecting the hypothesis that β_0 is equal across groups.

any subsector outside this range indicates the individual is outside the realm of plausibility of no consolidation problems. In 1997 one subsector, publishing and printing (NAICS 323), is outside the realm of plausibility.

Changes in the value of the percent differences from year to year are useful to determine if there is randomness in the OM and AR-1 data or if there are patterns to the differences. Though percent differences do bounce around between 2001 and 1997 for some subsectors, the majority are relatively stable. By far the largest change from 1997 to 2001 is publishing and printing which is 231.34% in 1997 and 11.14% in 2001. This is evidence of a data mistake (probably in AR-1) rather than a systematic problem with the subsector. No other subsector experiences so large a change in percent differences. The next largest are computers (334), electronic equipment (335), apparel (315), and leather (316).

Though miscellaneous manufacturing (339) has large differences, it is roughly constant over time. It is above 50% for both 2001 and 1997 and 60% in 1991. Though large, it is not unreasonable, and poses no real concern as it is a smaller subsector and a non-homogenous subsector since it composed of goods that do not fit into other categories.

Using the same analysis, look at individual states in table 3. Negative or small positive percent differences suggest the OM series underestimates the value of exports *above and beyond that attributed to f.o.b. vs. f.a.s.* Likewise, large positive values indicate the OM series overestimates the value of exports above and beyond that due to transportation costs.

Certain states stick out from the rest in table 3. Wyoming consistently has large positive percent differences. In all years except 2001, Wyoming has the largest percent difference. There are large differences between the data sets for Delaware, Florida, South Dakota, and Texas too. OM data indicates more than twice as much exports from Delaware as the AR-1. Florida and Texas have substantially increased their percent difference between the OM and AR-1 data in the most recent years. These states are far greater than 19.4% (10.55% in 1997) which is the percent difference for the United States. South Dakota reports the minimum percent difference in 2001, and is extremely low in every year.

A state is of poor quality if its percent difference, listed in table 3, is greater than 75% or less than -45% for a number years. Therefore a state with twice as much f.a.s. export value than f.o.b.

export values for three years is poor. Arkansas and Wyoming fall outside this range for all four years. For three years, the following states fall outside the range of acceptable quality: Florida, New Mexico, and South Dakota. Delaware, Hawaii, Idaho, Nevada, Texas, and Vermont are poor in two years. Though many of these states are small, several major states are listed such as Florida and Texas.

Scan over the values in table 3 by year. There is persistence for most states. Change occurs over time, but these changes are gradual. Virginia is an example. As the years progress, Virginia's percent difference gets smaller and becomes negative in 2001. Illinois takes the opposite path, beginning with a negative value and ending with a positive one. In fact, there is a trend for the Appalachian states of Tennessee, Kentucky, North Carolina, and West Virginia to gradually get positive differences. Of course, some states, small ones in particular, do have large swings in their percent differences. Vermont and Hawaii are examples. States with at least one wild fluctuation from year to year are Alaska, Colorado, Hawaii, Mississippi, Montana, and Vermont.

The descriptive test suggests printing and publishing (NAICS 323) is of low quality in at least one year and the OM data for Arkansas, Delaware, Florida, Hawaii, New Mexico, South Dakota, Texas, Vermont, and Wyoming do not represent the AR-1 data for these states.

4.5 Quality of the OM Data in the Early Years

Two studies compare the OM data to the AR-1 data for the initial years when both are available. These studies are Coughlin and Mandelbaum (1991) and Cronovich and Gazel (1999).[15] Despite the greater level of errors in the OM data in the early years, these studies agree with the findings presented here. Namely, the OM data is, overall, of good enough quality to use as the origin of production of exports.

Coughlin and Mandelbaum (1991; CM) compare data in one year only, 1987, the first year the OM series was published. CM compare ranks and also aggregate over all destinations and

[15]There are two other studies that are no longer available. Michael Farrell and Anthony Radspieler worked on an unpublished Census manuscript titled, "Census Bureau State-By-State Foreign Trade Data: Historical Perspectives; Current Situation; Future Outlook" in May 1989 and Michael Risha released a Manufacturing and Construction Division (Census) working paper, "A Comparison of the 'Origin of Movement' Series and the 'Exports from Manufacturing Establishments' Series" in 1991. This paper is not available for the public because it does not meet Census disclosure requirements.

manufacturing industries to consider total manufacturing exports by state. Differences in rank exist, but are small. Percent differences are widespread. Table 3 on page 18 reports their findings. CM fail to offer criteria for establishing if this variance is reasonable or not. They do not report a subsector detail, thus making their results difficult to interpret. Nonetheless, CM agree the OM data is similar to the AR-1 data for most states but warn of individuals of low quality.

Cronovich and Gazel (1999; CG), compare the OM to AR-1 series for the years 1987–1991. CG perform a number of cross sectional correlation exercises as well as fixed effects and non-fixed effects regressions. As mentioned, correlation is not a good indicator of agreement between data sets, but regressions are informative.

Using a regression pooled across all years, they find the OM data is not substitutable for AR-1 data in New Mexico, Tennessee, Vermont, Montana, New Hampshire and Wyoming. They issue low quality warnings for leather products (SIC 31, NAICS 316), petroleum and coal products (29, 324), textile mill products (23, 314), apparel (23, 315) and primary metals manufacturing (33, 331).

These two previous studies, using older data, arrive at the same conclusion as this present study: the OM data does have a problem with consolidation, but the quality is, in general, good enough to use for OP. Individual states or subsectors range from good enough to poor. Depending on the use, OM data for specific state or subsectors is not appropriate to use for OP.

5 Discrepancy Patterns and Modifications to Improve Quality

The comparison of the OM and AR-1 data shows the OM data is acceptable to use for the state of production of exports, but that some subsectors and states are of low quality. If, however, there are systematic over- or underestimates by state either for all manufacturing subsectors or only for a single subsector, then one may modify the data to correct for problems, and construct higher quality values. The tasks are to determine which states are consistently problematic and to determine if the reason for the problem is the same every year.

Since the percent differences in table 3 are largely persistent, they are likely based on state characteristics rather than randomness. The most obvious characteristic in this environment is location within the United States. Figure 1 codes states by percent differences in four separate

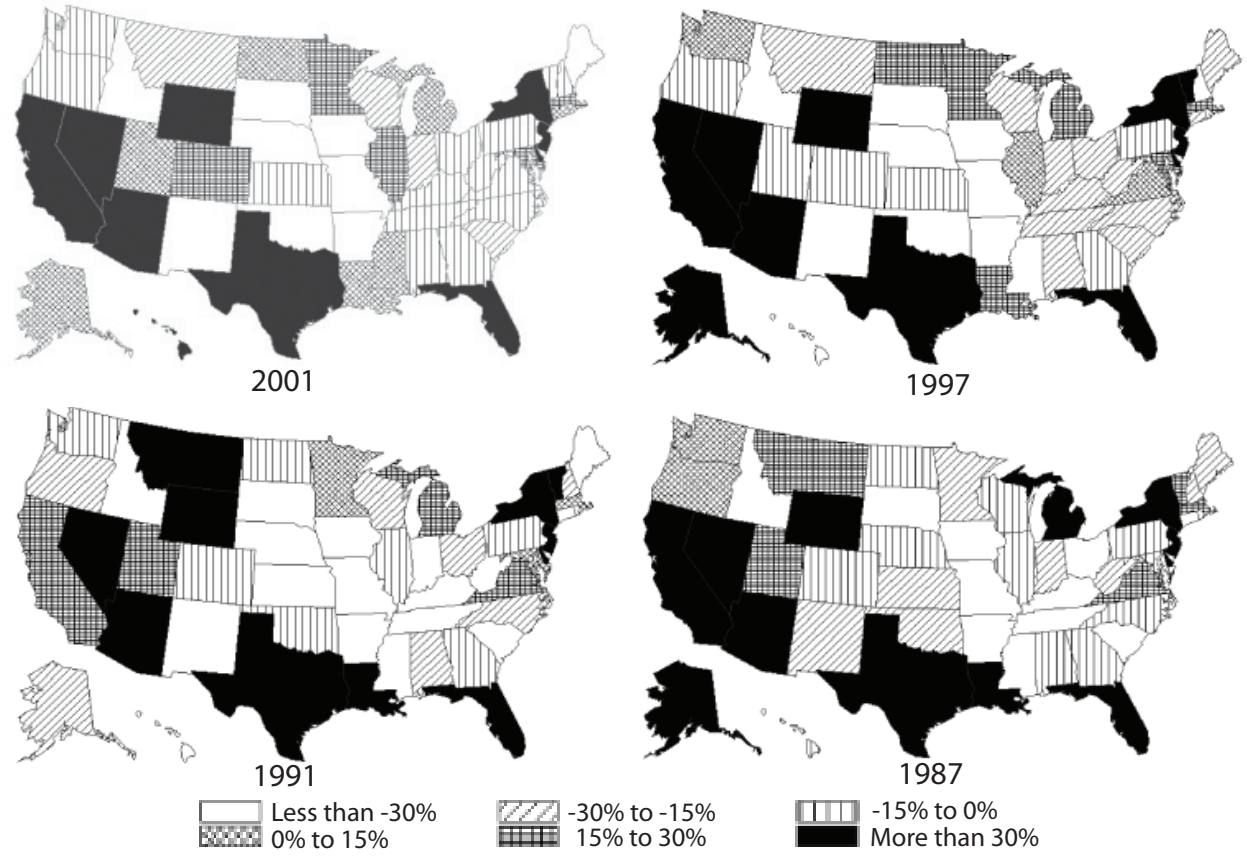

Figure 1. Percent differences of OM to AR-1 by state. Percent difference is $(OM - AR1)/AR1 * 100$; Author's calculations using OM data from WISER and AR-1 data from *Exports From Manufacturing Establishments*, Census.

images, one for each year. States coded in the two darkest colors have percent differences above 15% indicating the OM data over-estimates exports produced in those states. Border states such as California, New York, Florida, and Texas, are overestimated by OM year after year. Interior states such as South Dakota, Iowa, and Missouri are consistently underestimated by OM despite f.a.s. measurements. This pattern, however, does not hold strictly. Surprises include Rocky Mountain states. Nevada and Wyoming are consistently overestimated by OM despite their interior location. There are some surprises in the border states such as Washington, which is underestimated by OM. Even when the sign is as expected as in Ohio and Washington, the magnitude is much smaller than other port states.

Tables 8 for 2001 and table 9 for 1997 pull out port states from the rest using data on exports by port and by state location of ports from WISER and the SED. A state is a "port state" if the state exports more than 3% of total U.S. exports in all commodities—which due to data limitations

includes agriculture and mining. Hence, port states are: California, Texas, New York, Washington, Michigan, Illinois, Florida, and Ohio. Though it does not reach the 3% threshold, Louisiana is also included as a port state because the Port of New Orleans is so large. Table 16 in appendix C lists the fifteen states with the largest exports by ports. Table 17 shows the largest ports, by value, in the United States.

Not only is there potential for consolidation at ports, consolidation in foreign trade zones (FTZ) occurs as well. The SED instructs the USPPI to indicate the state of the FTZ rather than the state of production if the export travels through a FTZ. There are two types of FTZs: general and subzones. Subzones locate in the site of production and thus cause no problem. General subzones are near ports of exit and contribute to consolidation problems in the OM data. Table 18 in appendix C lists the largest general FTZs by exports in 2001. Though general FTZs may add to the consolidation problem, the effect is minimal since they account for only $3.8 billion (0.6%) of exports. Furthermore, since general FTZs are located in ports, most of the effect of FTZ consolidation will be in port states where consolidation is a problem regardless of the FTZ.

As seen in table 8, port states tend to have large positive percent differences between OM and AR-1. Washington and Ohio are the exceptions. The std. dev. of port states is 62.20. It is large because of Washington and Ohio. Non-port states ("small" states as well as "other" states) have a std. dev. of 45.09. Table 9 is similar. Therefore the large differences of the port states greatly contributes to the overall std. dev. in the state detail. The concern that the f.a.s. measurement in the OM data would inflate the values of exports of interior states due to greater transportation costs is, if anything, mitigating the effects of ship consolidation decreasing the value of interior states's exports.

Clearly the interior/port distinction goes a long way describing the patterns in the percent differences, but it is not enough. State size is another characteristic where patterns arise. One reason small states matter is observations from them are more likely to be non-disclosed than large states. This cannot explain tables 8 and 9, however, because state totals are used rather than the sum of reported subsectors.

A second reason why small states are more likely to have large differences is mistakes in data collection are less likely to correct themselves through large numbers. For example suppose Cali-

Table 8. Comparison of OM to AR-1 by port state and small state, 2001

State	Port			Small			Other		
	OM (Billions of $2001)	AR1 (Billions of $2001)	Difference (%)	OM (Billions of $2001)	AR1 (Billions of $2001)	Difference (%)	OM (Billions of $2001)	AR1 (Billions of $2001)	Difference (%)
AK	–	–	–	0.59	0.52	12.13	–	–	–
AL	–	–	–	–	–	–	6.86	7.56	−9.22
AR	–	–	–	–	–	–	2.79	5.16	−45.95
AZ	–	–	–	–	–	–	11.75	7.95	47.78
CA	98.43	70.32	39.98	–	–	–	–	–	
CO	–	–	–	–	–	–	5.90	4.90	20.40
CT	–	–	–	–	–	–	8.18	7.60	7.61
DE	–	–	–	1.89	0.81	134.98	–	–	–
FL	25.51	9.21	177.06	–	–	–	–	–	–
GA	–	–	–	–	–	–	13.50	14.42	−6.42
HI	–	–	–	0.30	0.17	73.97	–	–	–
IA	–	–	–	–	–	–	4.38	7.27	−39.76
ID	–	–	–	1.96	3.11	−36.88	–	–	–
IL	29.43	23.19	26.91	–	–	–	–	–	–
IN	–	–	–	–	–	–	14.05	17.76	−20.87
KS	–	–	–	–	–	–	4.53	5.27	−14.04
KY	–	–	–	–	–	–	8.43	8.53	−1.19
LA	8.81	7.87	11.86	–	–	–	–	–	–
MA	–	–	–	–	–	–	16.61	14.00	18.64
MD	–	–	–	–	–	–	4.62	4.01	15.12
ME	*v*	–	–	1.29	1.92	−32.64	–	–	–
MI	31.40	27.41	14.57	–	–	–	–	–	–
MN	–	–	–	–	–	–	9.73	8.25	17.99
MO	–	–	–	–	–	–	5.96	9.13	−34.72
MS	–	–	–	–	–	–	3.42	3.16	8.09
MT	–	–	–	0.29	0.40	−25.47	–	–	–
NC	–	–	–	–	–	–	15.73	17.61	−10.69
ND	–	–	–	0.63	0.60	5.69	–	–	–
NE	–	–	–	–	–	–	2.27	3.24	−30.10
NH	–	–	–	–	–	–	2.27	2.56	−11.36
NJ	–	–	–	–	–	–	17.56	10.63	65.22
NM	–	–	–	–	–	–	1.35	2.62	−48.26
NV	–	–	–	1.32	1.00	32.85	–	–	–
NY	36.40	21.13	72.27	–	–	–	–	–	–
OH	26.24	29.78	−11.89	–	–	–	–	–	–
OK	–	–	–	–	–	–	2.54	4.40	−42.20
OR	–	–	–	–	–	–	7.50	8.12	−7.55
PA	–	–	–	–	–	–	16.44	17.62	−6.69
RI	–	–	–	1.10	1.36	−19.18	–	–	–
SC	–	–	–	–	–	–	9.76	11.99	−18.59
SD	–	–	–	0.55	1.40	−60.56	–	–	–
TN	–	–	–	–	–	–	10.44	10.73	−2.70
TX	90.57	42.64	112.38	–	–	–	–	–	–
UT	–	–	–	–	–	–	3.31	3.30	0.27
VA	–	–	–	–	–	–	9.88	10.10	−2.20
VT	–	–	–	2.70	30.6	−11.64	–	–	–
WA	31.05	33.43	−7.12	–	–	–	–	–	–
WI	–	–	–	–	–	–	9.87	11.91	−17.09
WV	–	–	–	1.80	2.10	−14.09	–	–	–
WY	–	–	–	0.43	0.16	171.40	–	–	–
U.S.	622.36	521.39	19.37	622.36	521.39	19.37	622.36	521.39	19.37
Mean	41.98	29.44	48.45	1.14	1.28	17.74	8.20	8.56	−6.02
Std. dev.	–	–	62.20	–	–	69.33	–	–	26.27

Sources: OM data from WISER; AR-1 data from *Exports From Manufacturing Establishments*, Census.

Note: Percent difference is $(OM - AR1)/AR1 * 100$.

Table 9. Comparison of OM to AR-1 by port state and small state, 1997

State	Port			Small			Other		
	OM (Billions of $1997)	AR1 (Billions of $1997)	Difference (%)	OM (Billions of $1997)	AR1 (Billions of $1997)	Difference (%)	OM (Billions of $1997)	AR1 (Billions of $1997)	Difference (%)
AK	–	–	–	0.51	0.37	34.97	–	–	–
AL	–	–	–	–	–	–	5.40	7.24	−25.42
AR	–	–	–	–	–	–	2.16	4.23	−48.98
AZ	–	–	–	–	–	–	13.20	6.85	92.64
CA	91.29	67.84	34.56	–	–	–	–	–	–
CO	–	–	–	–	–	–	4.88	5.54	−11.88
CT	–	–	–	–	–	–	6.58	9.16	−19.38
DE	–	–	–	1.98	0.91	117.35	–	–	–
FL	21.86	9.99	118.78	–	–	–	–	–	–
GA	–	–	–	–	–	–	12.12	13.17	−8.00
HI	–	–	–	0.26	2.89	−90.90	–	–	–
IA	–	–	–	–	–	–	4.65	7.57	−38.58
ID	–	–	–	1.55	4.60	−66.22	–	–	–
IL	25.42	23.70	7.28	–	–	–	–	–	–
IN	–	–	–	–	–	–	11.72	16.47	−28.83
KS	–	–	–	–	–	–	3.91	4.06	−3.71
KY	–	–	–	–	–	–	7.50	10.20	−26.46
LA	9.28	7.53	23.23	–	–	–	–	–	–
MA	–	–	–	–	–	–	15.59	13.18	18.29
MD	–	–	–	–	–	–	4.97	4.31	15.44
ME	–	–	–	1.36	1.72	−21.11	–	–	–
MI	31.34	25.14	24.67	–	–	–	–	–	–
MN	–	–	–	–	–	–	8.52	7.40	15.17
MO	–	–	–	–	–	–	6.41	12.09	−47.01
MS	–	–	–	–	–	–	2.17	4.71	−53.93
MT	–	–	–	0.37	0.49	−23.98	–	–	–
NC	–	–	–	–	–	–	14.88	17.51	−15.02
ND	–	–	–	0.64	0.54	17.96	–	–	–
NE	–	–	–	–	–	–	1.77	2.94	−39.88
NH	–	–	–	–	–	–	1.47	3.22	−54.37
NJ	–	–	–	–	–	–	14.02	8.95	56.75
NM	–	–	–	–	–	–	1.73	6.21	−72.12
NV	–	–	–	0.85	0.61	40.13	–	–	–
NY	33.93	22.49	50.86	–	–	–	–	–	–
OH	23.85	29.49	−19.139	–	–	–	–	–	–
OK	–	–	–	–	–	–	2.63	4.03	−34.66
OR	–	–	–	–	–	–	7.27	7.43	−2.14
PA	–	–	–	–	–	–	15.39	16.76	−8.19
RI	–	–	–	0.92	1.53	−39.83	–	–	–
SC	–	–	–	–	–	–	7.27	9.26	−21.45
SD	–	–	–	0.46	1.04	−56.02	–	–	–
TN	–	–	–	–	–	–	8.71	10.30	−15.39
TX	71.43	39.08	82.75	–	–	–	–	–	–
UT	–	–	–	–	–	–	2.81	3.03	−7.08
VA	–	–	–	–	–	–	11.00	10.04	9.49
VT	–	–	–	3.74	1.33	181.77	–	–	–
WA	27.73	25.91	7.02	–	–	–	–	–	–
WI	–	–	–	–	–	–	9.34	11.60	−19.47
WV	–	–	–	1.42	1.78	−20.53	–	–	–
WY	–	–	–	0.52	0.04	1347.69	–	–	–
US	558.79	505.48	10.55	558.79	505.48	10.55	558.79	505.48	10.55
Mean	37.35	27.91	36.67	1.12	1.37	109.33	7.43	8.44	−14.08
Std. dev.	–	–	42.21	–	–	379.65	–	–	34.05

Sources: OM data from WISER; AR-1 data from *Exports From Manufacturing Establishments*, Census.

Note: Percent difference is $(OM - AR1)/AR1 * 100$.

fornia exports 200 packages each worth $2 million and Delaware exports 20 packages each worth $2 million. Further suppose ±$1 million mistakes occur in 10% of packages in both states. Then California has mistakes on 20 packages and Delaware on 2. Delaware has a 50% chance of having a 5% error in data collection since two of the four combinations of mistakes cause a $2 million error on $40 million worth of exports. California has a 0.05% chance of having a 5% error.

A "small state" is a state with a manufacturing share of less than 0.5% for both 2001 and 1997. These states are: Alaska, Hawaii, Wyoming, Montana, North and South Dakota, Nevada, Vermont, Delaware, Idaho, Rhode Island, Maine, and West Virginia. Notice no port state is also a small state. See table 19 in appendix C for a list of the smallest manufacturing states.

Of the thirteen small states in table 8, six have positive values. They also have extreme values. South Dakota is the minimum. Wyoming is the maximum. The std. dev. of small states is 69.33. Non-small states (including port states) are 44.02.

States that are neither port states nor small states have percent differences that tend to be negative. Arizona is overestimated by OM despite the fact it is the eighteenth largest exporter. Other border states without large ports such as Minnesota have the expected positive sign. States on the Mississippi tend to have negative values.

Looking back at table 3, an unexpected fact is, save Hawaii, extreme values are stable over time. Therefore if the reason small states have large percent differences is due to data mistakes, these mistakes occur in the same direction year after year. Table 10 breaks down selected state by NAICS subsector to determine if one subsector drives the percent differences for the entire state or not. The states chosen are ones of interest such as Wyoming and Nevada, as well as some port states such as Florida, and some "other states" such as Iowa and Kentucky. A similar table for 1997 follows.

The break down by state and subsector shows the problems of nondisclosure and truncation in the AR-1 data. Small states are more likely to have these issues in the AR-1 data than larger states because of the small number of exporters. Disclosure edits account for 5.2%, 5.5%, and 15.0% of direct exports by state and subsector in the 2001, 1997, and 1991 AR-1 respectively. However AR-1 reports zero direct exports for 18.3% of entries in 2001, 2.5% in 1997, and 21.9% in 1991. Therefore there is a nondisclosed or zero observation for 247 observations (23.5%) in 2001, 85

Table 10. OM and AR-1 by selected state and subsector, 2001

State	Data	311	312	313	314	315	316	321	322	323	324	325	326	327	331	332	333	334	335	336	337	339
												(Millions of $2001)										
AR	AR1	759.7	0.0	[a]	11.8	[a]	7.6	142.2	281.0	14.1	[a]	431.9	126.3	15.0	321.4	455.7	371.5	186.5	423.3	513.4	38.5	111.7
	OM	640.6	27.0	32.4	3.7	23.8	1.9	15.0	216.2	75.7	50.7	390.2	74.6	12.3	111.0	151.8	253.9	110.0	176.6	375.6	15.7	31.1
AZ	AR1	43.4	[a]	0.0	4.8	[a]	0.0	[a]	29.2	33.5	0.0	49.1	47.2	27.2	362.3	531.2	272.1	3003.6	70.7	3155.6	4.4	303.6
	OM	152.5	3.9	40.5	16.5	50.3	11.5	14.5	189.7	27.2	43.6	213.7	497.9	49.2	165.3	594.6	801.5	5506.6	501.9	2512.9	15.3	338.2
DE	AR1	24.8	0.0	0.0	0.0	0.0	0.0	[a]	19.8	1.6	0.0	428.8	111.8	0.0	0.0	24.9	18.5	15.1	0.0	72.2	0.0	31.1
	OM	53.1	0.0	12.1	7.5	2.4	0.4	1.3	27.5	3.1	7.2	916.4	99.0	4.4	18.95	47.8	124.5	253.1	21.4	243.2	7.6	41.7
FL	AR1	387.8	56.5	[a]	53.3	96.7	35.2	24.0	406.7	117.8	3.6	1298.1	88.4	56.1	70.0	230.0	807.9	2846.3	335.5	1231.6	66.9	969.9
	OM	1055.0	78.8	534.9	78.1	703.6	110.0	146.2	760.1	145.8	41.8	2569.9	481.6	167.1	145.3	582.4	3129.0	8382.6	111.1	3649.9	130.6	587.8
IA	AR1	1727.6	[a]	0.0	5.0	30.2	0.0	21.3	31.0	41.7	0.0	483.4	163.6	6.2	231.7	264.6	2331.1	847.4	490.9	375.3	11.5	181.1
	OM	951.0	8.3	26.3	15.1	5.0	2.3	29.1	35.7	39.2	3.4	426.6	125.7	8.9	213.7	144.3	1034.3	409.5	389.6	321.6	37.1	151.5
KY	AR	300.6	95.7	23.1	47.3	110.3	33.3	31.9	115.5	61.5	26.6	1250.1	220.9	331.8	446.8	240.1	761.5	1132.2	453.5	2745.5	7.5	99.0
	OM	105.6	210.7	83.2	16.7	110.3	13.4	73.4	91.5	116.5	8.9	1341.2	223.4	495.7	218.1	302.1	747.1	1131.8	279.4	3036.4	33.5	94.0
NY	AR1	307.1	8.1	249.0	74.4	606.3	74.9	172.6	283.2	287.0	10.0	3794.1	380.4	403.5	859.0	483.6	4419.2	4451.7	773.4	2569.7	64.9	857.0
	OM	661.2	76.6	382.0	106.2	423.1	135.6	192.5	670.4	839.2	145.6	4406.7	702.7	494.7	3892.9	705.8	4823.6	6537.6	1100.4	4467.0	84.7	550.7
NV	AR1	11.8	0.0	0.0	0.0	0.0	0.0	[a]	0.0	19.8	0.0	60.6	5.1	0.8	87.7	30.8	92.2	166.2	0.0	79.9	1.4	348.9
	OM	7.4	15.5	2.0	3.0	9.6	7.9	1.3	4.2	15.5	0.2	52.4	35.5	6.3	31.8	59.3	82.3	457.4	26.6	263.5	2.3	239.5
OH	AR1	386.1	12.6	88.6	17.7	123.4	27.5	116.5	381.7	355.1	62.1	2581.2	1047.7	649.1	1552.5	1728.6	3487.1	1579.6	1434.3	13593.6	166.5	383.1
	OM	475.6	10.6	104.4	55.2	48.9	24.6	152.9	464.8	130.4	81.3	2588.3	1070.9	609.1	1010.8	1613.3	4050.1	1676.7	1038.9	10353.1	123.7	552.1
SD	AR1	126.3	0.0	0.0	0.0	0.0	0.0	0.3	0.0	16.1	0.0	0.0	8.0	5.3	0.0	36.4	114.3	983.2	0.0	29.0	0.9	17.4
	OM	153.8	0.1	5.0	3.5	0.7	1.4	0.2	13.4	4.5	0.3	13.1	7.0	2.5	23.6	7.1	63.5	192.1	16.7	22.3	0.3	21.8
TX	AR1	2258.6	24.2	54.0	35.8	396.9	113.3	22.6	340.8	187.2	2745.3	11617.1	678.5	259.4	1218.8	1404.0	4585.3	13490.5	563.4	2093.8	96.7	458.7
	OM	2594.4	89.2	1364.1	196.4	771.4	604.3	123.1	1194.3	268.3	3705.0	14600.4	2763.3	650.3	2089.3	3198.8	12821.1	25684.0	4816.8	11258.0	151.3	1625.3
WA	AR1	883.7	22.5	20.2	28.8	39.3	0.0	463.8	720.0	56.0	395.6	249.9	73.5	117.6	529.4	124.1	767.0	2436.2	114.7	26137.4	11.3	224.7
	OM	1276.1	28.0	18.1	34.1	41.1	11.0	359.6	882.2	41.1	486.2	548.7	140.7	65.1	364.5	169.6	979.6	2229.1	260.4	22740.1	25.4	344.9
WY	AR1	0.0	0.0	0.0	0.0	0.0	0.0	0.3	0.0	0.0	0.0	113.9	0.0	0.0	0.0	0.0	7.2	0.0	0.0	0.0	0.0	0.0
	OM	1.7	0.0	0.0	0.0	0.1	0.0	1.7	0.1	0.1	0.1	387.0	0.3	1.1	1.9	8.1	12.0	8.2	6.4	2.5	0.0	0.4

Sources: OM data from WISER; AR-1 data from *Expors From Manufacturing Establishments*, Census.

[a]Withheld by Census to avoid disclosing data for individual companies.

Table 11. OM and AR-1 by selected state and subsector, 1997

State	Data	311	312	313	314	315	316	321	322	323	324	325	326	327	331	332	333	334	335	336	337	339
												(Millions of $1997)										
AR	AR1	780.6	2.3	6.4	8.0	43.9	7.8	55.8	318.2	2.0	3.4	517.6	127.4	34.8	370.2	229.6	483.9	194.8	347.8	595.2	16.3	86.5
	OM	373.6	9.9	10.8	4.4	35.7	8.8	24.0	121.8	42.5	5.1	340.2	59.7	13.4	181.0	126.7	179.2	106.8	165.6	314.5	14.2	21.5
AZ	AR1	64.4	0.0	3.2	2.8	9.4	2.9	6.4	65.4	18.2	0.1	100.0	47.2	14.6	443.3	466.8	220.3	3163.3	97.8	2010.0	3.9	109.5
	OM	75.0	4.1	37.3	26.2	55.2	7.8	6.2	107.0	13.0	13.2	136.9	196.1	20.8	251.0	336.9	639.5	9020.3	366.9	1561.0	21.9	299.3
DE	AR1	83.7	[a]	10.0	[a]	[a]	0.2	0.2	22.1	0.1	2.5	304.6	78.1	11.4	14.3	6.2	17.7	155.5	25.4	97.8	2.9	74.7
	OM	55.8	0.0	39.0	2.8	3.7	0.4	0.3	20.7	1.9	4.4	989.1	154.6	10.0	130.8	21.2	62.2	73.9	31.0	313.9	2.5	61.6
FL	AR1	630.8	57.0	16.7	39.8	129.5	50.0	40.1	637.4	31.2	8.5	2092.2	147.4	116.4	137.4	227.9	763.7	1902.7	351.4	1931.7	38.9	640.6
	OM	995.2	73.6	220.1	73.4	1041.4	88.7	121.0	706.4	157.0	55.9	2999.6	475.3	152.5	318.5	510.3	2564.8	6356.2	1137.2	2667.1	129.6	1014.9
IA	AR1	1644.8	0.0	[a]	[a]	48.6	39.4	26.8	37.8	17.3	0.3	561.2	290.5	6.0	474.1	274.5	2680.4	397.1	489.1	452.5	14.6	102.0
	OM	974.2	32.2	2.8	13.2	3.1	7.3	22.3	43.8	39.7	1.6	430.9	165.3	13.8	194.3	124.8	1569.5	334.8	254.7	288.8	31.1	100.5
KY	AR1	640.2	443.2	14.5	15.6	171.6	90.9	35.9	105.3	29.3	32.0	1450.4	201.8	257.7	440.2	198.1	822.2	1106.1	471.3	3552.8	6.0	114.3
	OM	86.5	137.2	10.8	4.9	263.5	39.5	41.4	84.7	128.3	7.4	825.1	165.4	121.1	283.5	155.1	862.5	731.7	300.7	3077.7	21.8	101.3
NV	AR1	8.5	0.2	??	7.1	1.0	[a]	0.5	6.2	8.4	[a]	53.0	23.0	8.9	105.5	50.1	39.4	129.5	8.9	44.9	4.3	109.0
	OM	4.4	0.3	1.2	4.3	11.5	2.2	3.1	4.0	11.6	0.3	47.6	33.7	6.1	248.2	28.1	76.4	173.4	15.7	72.7	1.4	106.8
NY	AR1	269.6	22.9	202.8	105.7	849.9	72.5	135.0	322.6	122.8	7.1	4040.4	429.7	410.5	690.1	697.8	5766.3	4135.9	890.1	2288.9	74.2	957.6
	OM	491.2	87.4	432.2	80.4	427.8	165.0	176.3	604.0	563.4	63.0	3602.4	550.2	357.7	4668.5	613.9	5720.5	6246.4	978.8	4179.9	71.8	3851.6
OH	AR1	269.0	3.1	71.3	18.0	176.9	26.3	82.8	297.4	143.2	35.4	2345.6	1042.7	682.8	1426.0	1670.3	3828.8	1867.2	1434.8	13624.4	163.4	282.6
	OM	247.5	8.7	80.3	56.6	44.1	57.4	104.5	263.9	134.0	48.9	2713.3	883.9	592.5	944.3	1119.9	4380.4	1922.0	922.5	8852.2	104.3	369.2
SD	AR1	88.9	0.0	[a]	2.0	[a]	[a]	1.6	3.0	9.5	[a]	36.2	0.5	2.3	3.4	7.6	83.2	694.8	12.0	54.0	2.3	19.6
	OM	53.8	0.0	6.2	5.6	0.4	0.3	0.8	15.5	4.1	1.0	13.8	7.1	0.6	44.6	7.1	52.3	163.0	11.5	25.1	0.4	42.9
TX	AR1	1774.6	25.1	33.1	16.3	683.6	27.5	106.3	465.0	56.1	2673.2	11279.4	415.1	232.0	1336.3	1380.2	3844.0	10847.6	689.2	2690.1	71.2	438.6
	OM	1912.5	69.6	525.3	124.2	1013.1	456.2	133.8	983.8	198.2	2936.4	12375.7	1722.6	449.8	2066.8	2538.9	11276.8	19076.0	3740.9	8659.2	96.7	1069.5
WA	AR1	691.9	76.0	15.8	57.9	51.9	20.9	699.1	1058.1	45.5	229.2	310.1	137.8	69.7	889.9	144.5	625.4	2281.7	105.6	18208.5	20.4	172.9
	OM	1069.9	42.2	14.0	24.6	29.5	24.6	640.6	899.5	76.7	352.9	647.5	135.5	51.3	671.4	212.6	929.7	1658.3	238.0	19710.9	25.6	273.0
WY	AR1	3.7	0.0	[a]	[a]	[a]	[a]	[a]	0.0	0.0	0.9	2.7	0.9	0.5	0.3	4.9	14.5	3.5	[a]	2.1	0.0	0.1
	OM	1.4	0.0	0.0	0.0	0.2	0.0	1.0	0.1	1.2	0.0	488.9	0.7	0.2	2.2	4.3	11.1	8.7	1.0	2.7	0.0	0.4

Sources: OM data from WISER; AR-1 data from *Expors From Manufacturing Establishments*, Census.

[a]Withheld to avoid disclosing data for individual companies.

observations (8.1%) in 1997, and 369 (36.9%) in 1991. For most states, disclosure edits account for a small portion of the total reported for each state. For some states, however, disclosure and truncation edits account for the bulk of exports. In 2001 more than half of the manufacturing totals for the following states are not reported when disaggregated by NAICS subsector: Alaska (71%), Hawaii (54%), Idaho (86%), Montana (73%), and Vermont (82%). These states are not included in the select group of table 10. The problem of nondisclosure and truncation occurs for several subsectors too. In 2001 the sum over states for the following subsectors is less than 85% of the reported total: beverage and tobacco products (83%), apparel (80%), and petroleum and coal products (77%).

The evidence from table 10 and its equivalent for 1997, table 11, as well as figure 5 in appendix B, indicate that for some states, a single industry is creating large percent differences every year. For example, the large percent difference in Wyoming is due to several hundred million in chemicals exports in both 2001 and 1997. AR-1 reports zero exports for Wyoming chemicals in 1991 whereas OM reports $289 million. The same is true in Delaware where chemcials cause the large percent difference going back to 1991.

Nevada is different. In 2001 transportation equipment (336) makes the percent difference positive, but in 1997 the difference in the data sets is much smaller. In 1997 apparel (315) is the source. It is plastics and rubber (326) in 1991. This suggests the percent difference between OM and AR-1 in Nevada is due to random mistakes that for some reason turn up positive every year.

The large negative percent difference in South Dakota is due to computers and electronic equipment (334) in both 2001 and 1997. The subsector is not disclosed in 1991. Transportation equipment (336) is in large part responsible for the negative percent differences in Ohio and Washington, although other subsectors show negative percent differences also. For Ohio, the notable subsector is primary metal manufacturing (331). On the other hand, Washington has a large positive percent difference in chemicals (325) going back to 1991.

Florida and Texas are overestimated by OM in nearly every subsector each and every year. This indicates it is consolidation along the port rather than an idiosyncratic subsector.

The lessons from sections 4.4, 4.5, and 5, suggest time invariant consolidation problems are largest in chemicals (NAICS 325), machinery (333), computers and electronic products (334), and

transportation equipment (336). These are large subsectors that often cause low quality for states that otherwise would have high enough quality. Subsectors state pairs such as Wyoming chemicals and South Dakota computers should be removed from the data. Transportation equipment (336) plays a large role in the percent differences in the data for Ohio and Washington. Rather than modifying such observations, the recommended approach is to use the OM data but to be flexible about results obtained. Removing these predictably over- and underestimated observations improves the limits of agreement.

Besides consolidation, randomness may cause large percent differences. For example, printing and publishing (NAICS 323) is a small subsector where randomness is a problem. Coal and petroleum products (324) is another. One modification to small subsectors to avoid randomness is to combine them. Beverage and Tobacco products (NAICS 312) should combine with food products (311), textiles and fabric (313) and textile mill products (314) should join together, and apparel (315) and leather products (316) should join together. This procedure did *not* dramatically improve quality when tested.

For the same reason, small states such as Alaska, Delaware, Hawaii and Vermont are poor. Combining small states with larger neighbors is not a good solution. This is because state policy differs. Policy information is lost when states are grouped.

Consolidation at ports of exit does cause the OM data to overestimate exports by OP for the majority of ports states, though not Washington and Ohio. Port states are overestimated in every subsector. Interior states are underestimated, but frequently this is due to a single large subsector; not all subsectors. Small states greatly affect the overall variance. The cause is usually, but not always, random data collection mistakes. Some small states have consistent mistakes. The OM data may be used for the origin of production of manufactured exports, but caution is urged with the results particulary for Florida and Texas, and small interior states.

6 Conclusion

The OM series is the only data set with export values by state, subsector, and destination. There is concern about the quality of the data as it relates to the state of origin of movement for an export

Table 12. Subsectors and states with poor quality

OM and AR-1 not substitutable	NAICS	325, 333, 334, 336
	States	AR, DE, FL, NM, SD, TX, VT, WY
OM does not represent OP	NAICS	All non-manufacturing, 312, 313, 316, 323, 325, 333, 334, 336
	States	AK, AR, DE, FL, HI, NM, SD, TX, VT, WY

versus the state of production. To confirm the quality of the OM series, the collection of the data is described in detail to indicate possible errors. Limitations of the data are discussed. A battery of tests are performed to establish both the overall representativeness of the OM data to OP and the AR-1 data.

Results indicate the OM series data is inflicted by problems of consolidation at ports of exit, but nonetheless is of high enough quality to use for state manufacturing exports for the majority of states and subsectors. The results are consistent with tests performed by other authors for earlier years of the data, when electronic submission of export declariations was non-existent.

Errors in the OM data compared to the AR-1 data on origin of production of exports are largely the results of data collection mistakes in small states and subsectors and problems of consolidation at ports. Often the cause for a state to have questionable quality is due to a single, large subsector. Results from subsectors and states where consolidation leads to poor quality of OM data should be treated accordingly. Small subsectors and states do not pose a problem for using OM for OP since any data on OP will have problems with small subsectors and states.

The recommended modification is to eliminate non-manufacturing subsectors from consideration, and to be skeptical of results using printing and publishing (323), chemicals (325), machinery (333), computers and electronic products (334), and transportation equipment (336). To the extent small subsectors are of poor quality, the poor quality is due to their smallness rather than consolidation. The recommended modification to states is the removal of small states from consideration. This applies primarily to Alaska and Hawaii, though other candidates for removal are Delaware and Vermont. Florida and Texas clearly benefit from consolidation in the OM data whereas Arkansas and New Mexico are harmed. Data from these states should be treated accordingly. Finally particular state-subsectors pairs such as Wyoming chemicals and South Dakota computers should be eliminated from the OM data when using it for the origin of production of

exports.

References

Altman, D.G, and J.M. Bland 1983. "Meaurement in Medicine: The Analysis of Method Comparison Studies" *Statistician* 32 (3), 307–317.

Bernard, Andrew B., and J. Bradford Jensen 1995. "Exporters, Jobs, and Wages in U.S. Manufacturing: 1976–1987" *Brookings Papers on Economic Activity: Microeconomics* 1995, 67–112.

Bureau of the Census and Bureau of Economic Analysis, U.S. Department of Commerce various yearsa. *FT-900.* Washington, DC. `http://www.census.gov/ft900` (accessed April 10, 2006).

Bureau of the Census, U.S. Department of Commerce various yearsb. *Exports from Manufacturing Establishments.* Washington, DC. Analytic Report Series; AR(year)-1. Some data available as .pdf at `www.census.gov\mcd\exports`.

Coughlin, Cletus C., and Howard J. Wall 2003. "NAFTA and the Changing Pattern of State Exports" *Papers in Regional Science* 82 (4), 427–450.

Coughlin, Cletus C., and Patricia S. Pollard 2000. "State Exports and the Asian Crisis" *Federal Reserve Bank of St. Louis Review* 82 (1), 3–13.

——— 2001. "Comparing Manufacturing Export Growth Across States: What Accounts for the Differences?" *Federal Reserve Bank of St. Louis Review* 83 (1), 25–40.

Coughlin, Cletus C., and Thomas B. Mandelbaum 1991. "Measuring State Exports: Is There a Better Way?" *Federal Reserve Bank of St. Louis Review* 73 (4), 65–79.

Cronovich, Ron, and Ricardo Gazel 1999. "How Realiable Are the MISER Foreign Trade Data?" unpublished.

Feenstra, Robert C. 1997. "U.S. Exports, 1972-1994, With State Exports and Other U.S. Data" NBER working paper no. 5990. OM data available at `http://cid.econ.udavis.edu/usixd/wp5990d.html` (accessed April 20, 2006).

Feenstra, Robert C., John Romalis, and Peter K. Schott 2002. "U.S. Imports, Exports and Tariff Data, 1989-2001" NBER working paper no. 9387.

Foreign Trade Division, Bureau of the Census, U.S. Department of Commerce 1994. *U.S. Merchandise Trade Statistics, A Quality Profile.* Washington, DC. `http://www.census.gov/foreign-trade/misc/qprofile.htm` (accessed April 15, 2006).

Foreign Trade Division, Bureau of the Census, U.S. Department of Commerce 2002. *Correct Way to Complete the Shipper's Export Declaration Form 7525-V.* Washington, DC. `http://www.census.gov/foreign-trade/regulations/forms/correct-way-to-complete-the-sed.pdf` (accessed April 1, 2006).

Foreign Trade Division, Bureau of the Census, U.S. Department of Commerce 2005. *A Profile of U.S. Exporting Companies, 2002–2003.* Washington, DC. `http://www.census.gov/foreign-trade/aip/index.html` (accessed March 20, 2006).

Foreign Trade Zones Board, U.S. Deparment of the Treasury and U.S. Department of Commerce 2003. *63th Annual Report of the Foreign-Trade Zones Board to the Congress of the United States.* Washington, DC.

Kehoe, Timothy J., and Kim J. Ruhl 2004. "The North American Free Trade Agreement After Ten Years: Its Impact on Minnesota and a Comparison with Wisconsin" *Center for Urban and Regional Affairs Reporter* 34 (4), 1–10.

Lawson, Ann M, Kurt S. Bersani, Mahnaz Fahim-Nader, and Guo Jiemin 2002. "Benchmark Input-Output Accounts of the United States, 1997" Bureau of Economic Analysis, U.S. Department of Commerce. Washington, DC.

Minnesota Trade Office, Department of Employment and Economic Development various years. *Minnesota Quarterly Export Statistics.* St. Paul, MN. Contact: Christopher Sprung. `http://www.exportminnesota.com/itradestats.htm` (accessed May 1, 2006).

Office of Trade and Industry Information, International Trade Administration,U.S. Department of Commerce n.d. *Guide to State Export Data.* Washington, DC. `http://www.ita.doc.gov/td/industry/otea/state/technote.html` (accessed March 10, 2006).

Smith, Pamela J. 1999. "Are Weak Patent Rights A Barrier to U.S. Exports?" *Journal of International Economics* 48, 151–177.

Smith, Tim R. 1989. "Regional Exports of Manufactured Products" *Federal Reserve Bank of Kansas City Economic Review* 74 (1), 21–31.

Suits, Daniel B. 1984. "Dummy Variables: Mechanics V. Interpretation" *Review of Economics and Statistics* 66 (1), 177–180.

U.S. Government Printing Office 2005. *Commerce Control List.* Washington, DC. `http://www.access.gpo.gov/bis/index.html` (accessed June 20, 2006).

World Institute for Strategic Economic Research various years. *Origin of Movement State Export Data.* Holyoke, MA. `http://www.wisertrade.org` (accessed Nov. 15, 2005).

Wysocki, Adam J. 1998. "Outbound Process Improvements: Quality of U.S. Export Trade Statistics" Foreign Trade Division, Bureau of the Census, U.S. Department of Commerce. Washington, DC. `http://www.census.gov/foreign-trade/aip/outbound.pdf` (accessed April 22, 2006).

Appendices

In addition to the appendices below, there is an online appendix with some data (OM is proprietary), SAS code, and nonessential tables. This appendix is available at `www.econ.umn.edu/~cassey`.

A Other Sources of State Export Data

From 1992 through 2002, the Census published the Exporter Location (EL) data set which includes exports to foreign destinations by state and subsector. This data differs from the OM in that the former was based on the state of the ZIP code of the exporter's billing address. Thus EL was compiled using a different question on the SED than the OM. The billing address in practice is often the headquarters of a multi-plant firm or wholesaler. Thus EL, like OM, was not designed to represent the OP of exports. As with OM, raw EL data was adjusted by WISER until 2000.

The EL series was canceled in 2002 because of *monthly* wild swings in data from 1999–2002. These swings were caused by the increase in electronic submission of the SED. Since electronic filing reduces errors, the Census feels the EL data is flawed.

AR-2 data is published as *Selected Characteristics of Manufacturing and Wholesale Establishments That Export*. It was published every five years until 1992. The only years of overlap with the OM series are 1987 and 1992. The data comes from surveys from the Census of Manufactures and Census of Wholesale Trade, of the Economic Census. Like the AR-1, AR-2 data is adjusted with data from the SED.

AR-2 differs from AR-1 in that data is available, subject to disclosure concerns, for both manufacturing and wholesale firms with at least one paid employee. Thus one can determine the exports by wholesalers by state and subsector.

B Scatterplots of OM vs. AR-1

Figure 2 is a graphical representation of the relation of the OM and AR-1 data. The twenty-one manufacturing NAICS subsectors are plotted for 2001 and 1997. The left image is for the entire data set. The right image is a condensed image of the left image to highlight the smaller subsectors that are indistinguishable in the left image. The regression is re-run on the right so that the effects of large outlying points are removed.

In each image, the dashed line is the regression line. Because of the difference between f.a.s. and f.o.b. costs, one should not expect the regression to be the 45-degree line. Instead one expects the data to lie on a line with slope equal to the percent difference between f.a.s. and f.o.b. Fifteen percent is chosen, somewhat arbitrarily, based on the percent difference of the national total over a number of years. Hence the solid line has slope 1.15. Since each subsector has its own conversion between f.a.s. and f.o.b., care is needed in interpreting these figures. It is not necessarily true that a subsector above the solid line is overestimated by the OM series. It could be that subsector has greater transportation mark-up than other subsectors. The conversion for each subsector is not used so as to be consistent with state regressions forthcoming.

The linear association between AR-1 and OM is 1.27 in 2001. It is 1.22 in 1997. All slope coefficients in the following four figures are significantly different from zero at the 95% level .

Figure 3 is a graphical comparison of the OM and AR-1 data for the years 2001, 1997, 1991, and 1987. Moving left to right in each row highlights smaller states which are indistinguishable in the leftmost image. As with figure 2, each condensed image corresponds to a re-run regression without the outlying points. Also like figure 2, the dashed line is the linear association between the two data sets and the solid line represents the line of equality adjusted for transportation costs of 15%.

The linear association between the OM and AR-1 data is 1.43 in 2001 and 1.36 in 1997. This compares favorably to 1.29 and 1.26 for the subsector detail in 2001 and 1997 respectively. Again, overall the OM data is of high enough quality to use as the origin of production by state.

Notice in the leftmost images, the linear association has greater slope than the solid line. The opposite is true in the right most images. This indicates OM attributes more exports to larger states than to smaller states than it should based on the distinction of f.a.s. to f.o.b. Furthermore, the intercepts are negative in the leftmost images, but positive in the right most. This indicates smaller states are under-estimated by OM, but the smallest states are over-estimated. Of the smallest states, it is interior states such as Arkansas responsible for the flat slope. This corroborates the finding of consolidation at port states.

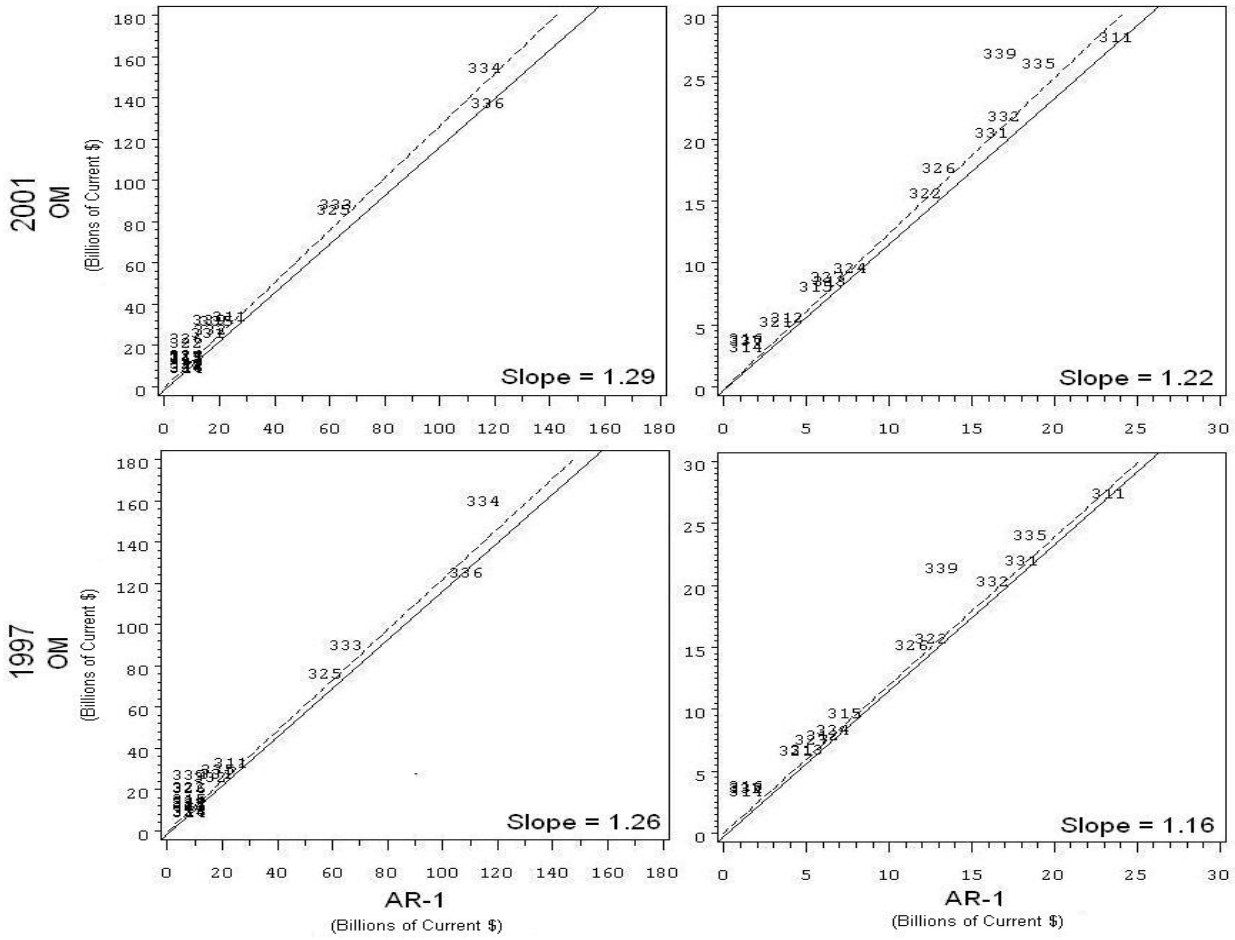

Figure 2. OM vs. AR-1 by subsector. The dashed line is the linear association between AR-1 and OM. The solid line has slope 1.15. It represents the linear association when the percent difference between f.a.s. and f.o.b. is 15%. The OM data is from WISER; the AR-1 data is from *Exports From Manufacturing Establishments*, Census.

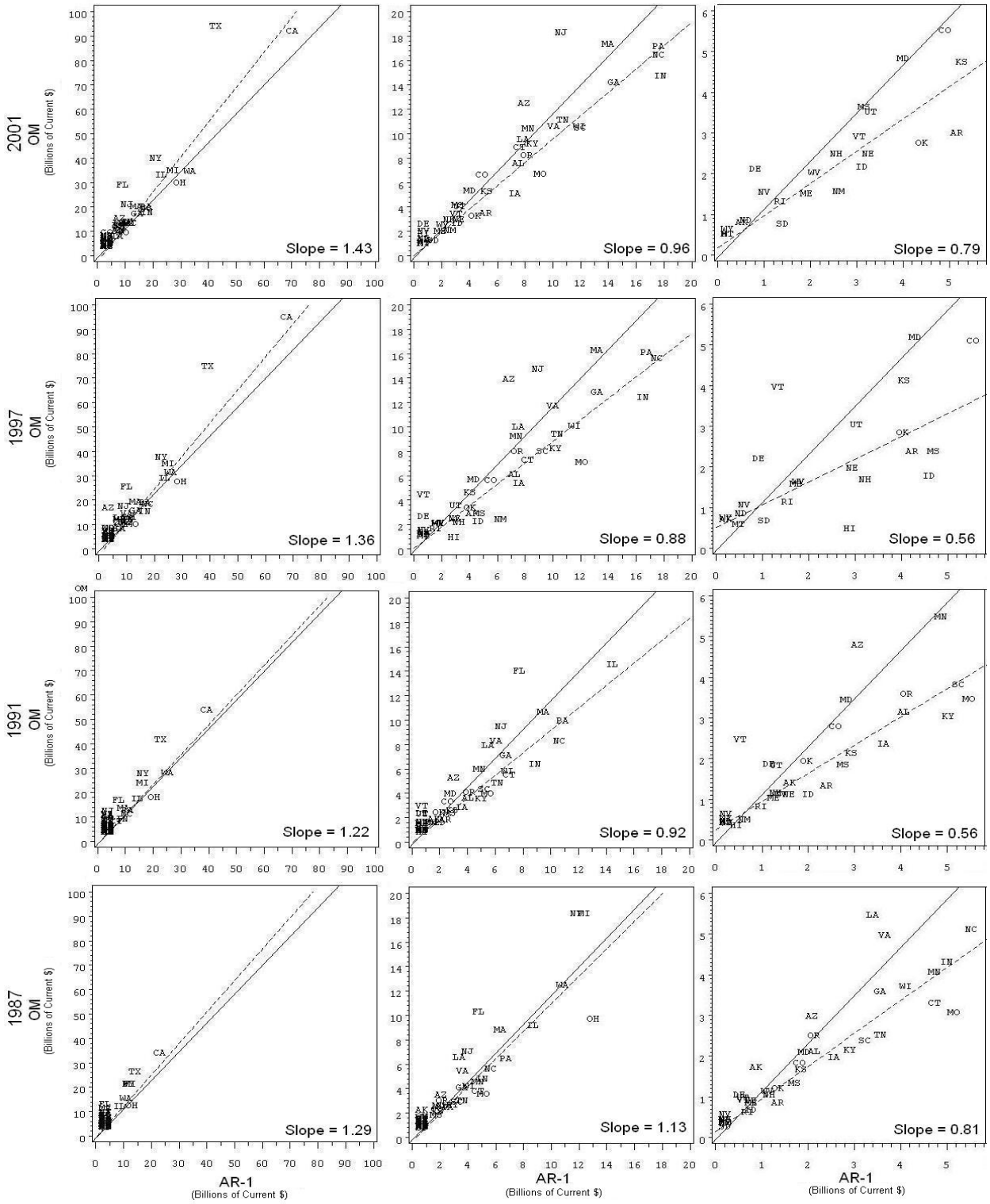

Figure 3. OM vs. AR-1 by state. The dashed line is the linear association between AR-1 and OM. The solid line has slope 1.15. It represents the linear association when the percent difference between f.a.s. and f.o.b. is 15%. The OM data is from WISER; the AR-1 data is from *Exports From Manufacturing Establishments*, Census.

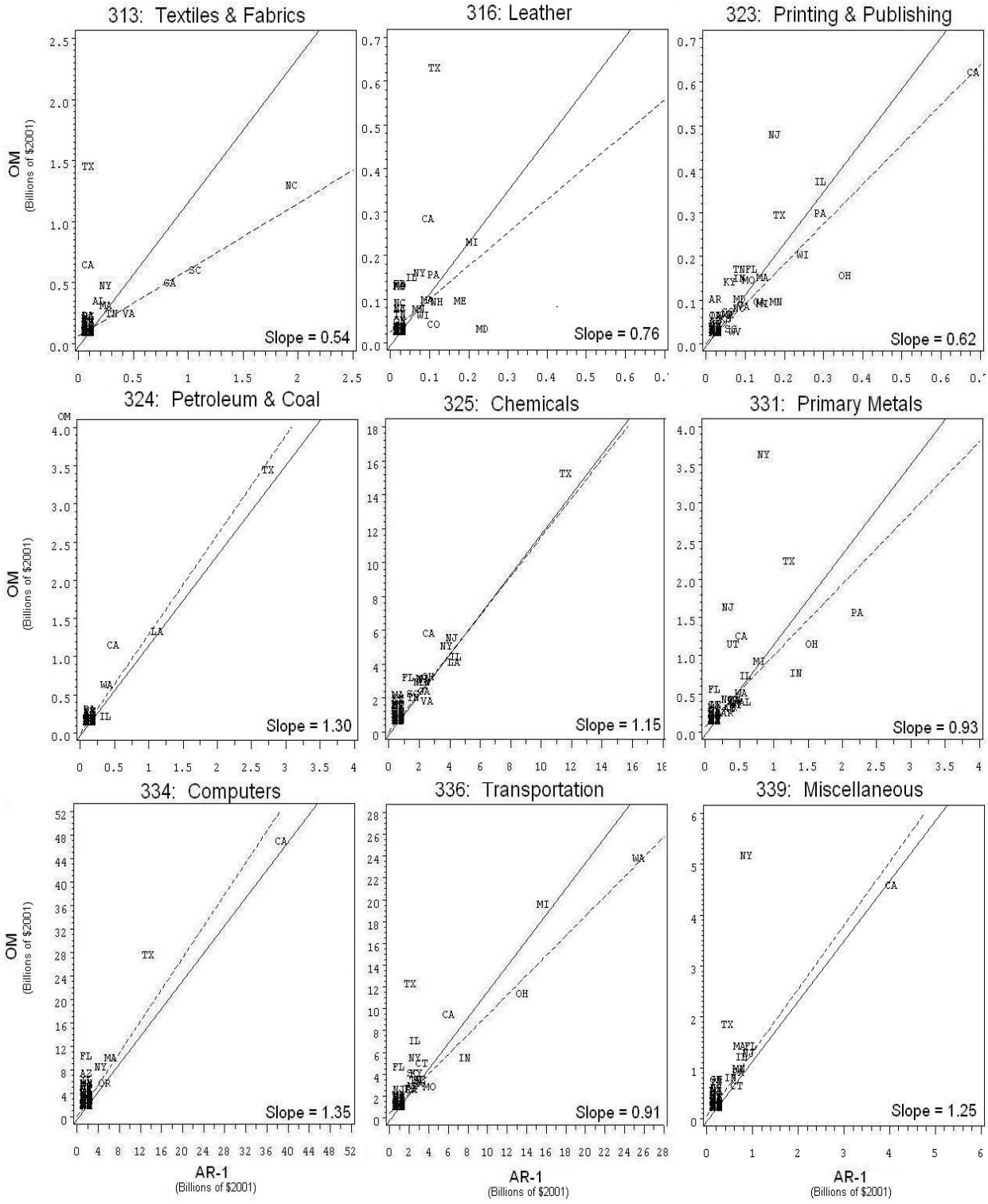

Figure 4. Linear association of selected subsectors, 2001. The dashed line is the linear association between AR-1 and OM. The solid line has slope 1.15. It represents the linear association when the percent difference between f.a.s. and f.o.b. is 15%. The OM data is from WISER; the AR-1 data is from *Exports From Manufacturing Establishments*, Census.

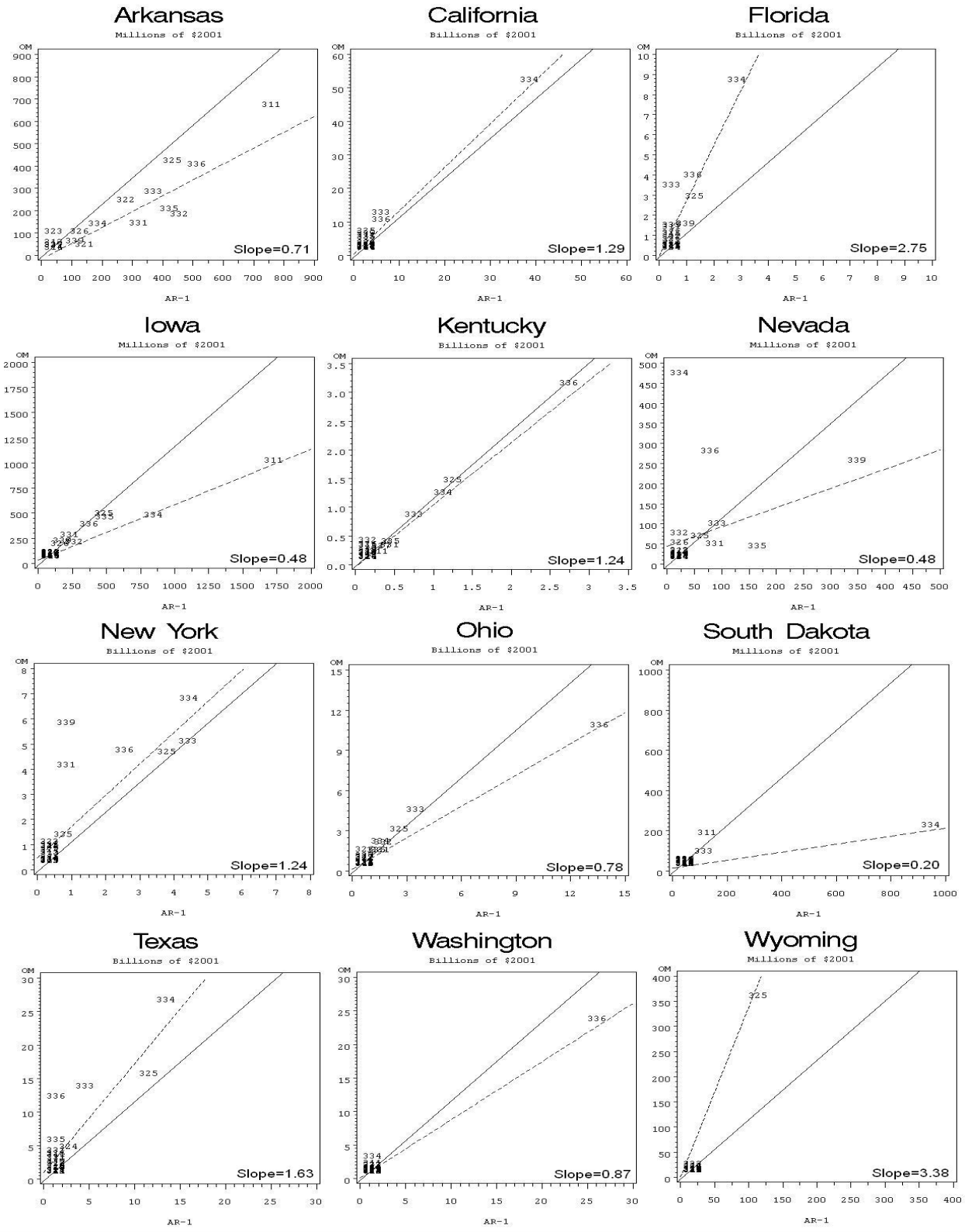

Figure 5. Linear association of selected states, 2001. The dashed line is the linear association between AR-1 and OM. The solid line has slope 1.15. It represents the linear association when the percent difference between f.a.s. and f.o.b. is 15%. The OM data is from WISER; the AR-1 data is from *Exports From Manufacturing Establishments*, Census.

C Supplemental Tables

Table 13. NAICS (1997-present) and SIC (1987-2000) subsectors in the OM series

NAICS	SIC	AR-1	Description
111	01		Agricultural products
112	02		Livestock and livestock products
113	08		Forestry products, NESOI[a]
114	09		Fish, fresh, chilled or frozen and other marine products
211	13		Oil and gas
212	10,12,14		Minerals and ores
311	20	Y	Food and kindred products
312	20,21	Y	Beverages and tobacco products
313	22	Y	Textiles and fabric
314	22,23	Y	Textile mill products
315	22,23	Y	Apparel and accessories
316	31	Y	Leather and allied products
321	24	Y	Wood products
322	26	Y	paper
323	27	Y	Printing, publishing, and similar products
324	29	Y	Petroleum and coal products
325	28	Y	Chemical
326	30	Y	Plastics and rubber products
327	32	Y	Nonmettalic mineral products
331	33	Y	Primary metal manufacturing
332	34,35	Y	Fabricated metal products, NESOI[a]
333	35	Y	Machinery, except electrical
334	35,36,38	Y	Computer and electronic products
335	36	Y	Electrical equipment, applicances, and components
336	37	Y	Transportation equipment
337	25	Y	Furniture and fixtures
339	38,39	Y	Miscellaneous manufactured commodities
511	27		Prepackaged software[b]
910	91		Waste and scrap[c]
920	92		Used or second-hand merchandise[c]
980	93		Goods returned to Canada[c]
990	95,99		Special classication provisions, NESOI[a,c]

Source: Descriptions are from WISER who adjusts descriptions to reflect export goods as in schedule B, not all commodities in the category.

Notes: Y under the AR-1 column denotes the subsector is available in AR-1, subject to disclosure concerns; Since concordance between 2002 NAICS and 1987 SIC at this level of aggregation is not exact, descriptions for NAICS and SIC may differ slightly.

[a]NESOI is a standard acronym for "not elewhere specified or included."

[b]511 is Publishing Industries which is mostly non-tradeables. However, 511 does contain one tradeable good, prepackaged software.

[c]These represent schedule B codes that do not have NAICS or SIC classifications. The Foreign Trade Division of the Census has created "trade-related" NAICS and SIC to handle such goods. These codes do not appear in a regular list of NAICS or SIC published by the Census.

Table 14. Foreign destinations 1987-present

Destination	ISO	U.S.	UN
Afghanistan	AFG	5310	450040
Albania	ALB	4810	580080
Algeria	DZA	7210	130120
American Somoa	ASM	9510	728882
Andorra	AND	4271	532500
Angola	AGO	7620	160240
Anguilla	AIA	2481	356580
Anitgua	ATG	2484	356580
Argentina	ARG	3570	330320
Armenia	ARM	4631	460510
Aruba	ABW	2779	355320
Australia	AUS	6021	710360
Austria	AUT	4330	550400
Azerbaijan	AZE	4632	460310
Bahamas	BHS	2360	350440
Bahrain	BHR	5250	440480
Bangladesh	BGD	5380	450500
Barbados	BRB	2720	350520
Belarus	BLR	4622	581120
Belgium	BEL	4231	530560
Belize	BLZ	2080	360840
Benin	BEN	7610	162040
Bermuna	BMU	2320	220600
Bhutan	BTN	5682	450000
Bolivia	BOL	3350	330680
Bonsia-Hercegovina	BIH	4793	590700
Botswana	BWA	7930	117100
Brazil	BRA	3510	330760
British Indean Ocean Territory	IOT	7810	166900
Brunei	BRN	5610	450000
Bulgaria	BGR	4870	581000
Burkina	BFA	7600	168540
Burma (Myanmar)	MMR	5460	451040
Burundi	BDI	7670	161080
Cambodia	KHM	5550	451160
Cameroon	CMR	7420	141200
Canada	CAN	1220	211240
Cape Verde	CPV	7643	166240
Cayman Islands	CYM	2440	353880
Central African Republic	CAF	7540	141400
Chad	TCD	7560	141480
Chile	CHL	3370	331520
China (Mainland)	CHN	5700	481560
China (Taiwan)	TWN	5830	458960
Christmas Island	CXR	6024	710360
Cocos Islands	CCK	6023	710360
Colombia	COL	3010	331700
Comoros	COM	7890	166380
Congo (Congo-Brazzaville)	COG	7630	141780
Cook Islands	COK	6142	715540
Costa Rica	CRI	2230	341880
Cote d'ivoire	CIV	7480	163840
Croatia	HRV	4791	591910
Cuba	CBA	2390	351920
Cyprus	CYP	4910	441960
Czech Republic	CZE	4351	582030
Czechoslovakia	CSK	4350	582000
Denmark	DNK	4099	532080
Djibouti	DJI	7770	162620
Dominica	DMA	2486	356580
Dominican Republic	DOM	2470	352140
East Timor	TLS	5601	453600
Ecuador	ECU	3310	332180
Egypt	EGY	7290	138180
El Salvador	SLV	2110	342220
Equatorial Guinea	GNQ	7380	162260
Eritrea	ERI	7741	162300
Estonia	EST	4470	582330
Ethiopia (old)	ETH	7740	162300
Ethipia (new)	ETH	7749	162300
Falkland Islands	FLK	3720	362380
Faroe Island	FRO	4091	532080
Fiji	FJI	6863	722420
Finland	FIN	4050	552460
France	FRA	4279	532500
French Guiana	GUF	3170	362540
French Polynesia	PYF	6414	725400
French Southern and Antarctic Lands	ATF	7905	166380
Gabon	GAB	7550	142660
Gambia, The	GMB	7500	162700
Gaza Strip (Israel)	PSE	5082	413760
Georgia	GEO	4633	462680
Germany, East	—	4290	582780
Germany, Federal Republic of	DEU	4280	532800
Ghana	GHA	7490	162880
Gibraltar	GIB	4720	572920
Greece	GRC	4840	533000
Greenland	GRL	1010	223040
Grenada	GRD	2489	356580
Guadeloupe	GLP	2831	353120
Guam	GUM	9350	368960
Guatemala	GTM	2050	343200
Guinea	GIN	7460	163240
Guinea-Bissau	GNB	7642	166240
Guyana	GUY	3120	363280
Haiti	HTI	2450	353320
Heard and McDonald Islands	HMD	6029	710360
Hondurus	HND	2150	343400
Hong Kong	HKG	5820	453440
Hungary	HUN	4370	583480
Iceland	ISL	4000	553520
India	IND	5330	453560
Indonesia	IDN	5600	453600
International Organizations	—	8500	—
Iran	IRN	5070	443640
Iraq	IRQ	5050	443680
Iraq-Saudi Neural Zone	NTZ	5160	446820
Ireland	IRL	4190	533720
Israel (new)	ISR	5081	413760
Israel (old)	ISR	5080	413760
Italy	ITA	4759	533800
Jamaica	JAM	2410	353880
Japan	JPN	5880	413920
Jordan	JOR	5110	444000
Kazakstan	KAZ	4634	463980
Kenya	KEN	7790	164040
Kiribati	KIR	6226	722960
Korea, Republic of	KOR	5800	454100
Kuwait	KWT	5130	444140
Kyrgyzstan	KGZ	4635	464170
Laos	LAO	5530	454180
Latvia	LVA	4490	584280
Lebanon	LBN	5040	444220
Lesothto	LSO	7990	117100

Table 14 continued

Destination	ISO	U.S.	UN
Liberia	LIB	7650	164300
Libya	LBY	7250	134340
Liechtenstein	LIE	4411	557560
Lithuania	LTU	4510	584400
Luxembourg	LUX	4239	530560
Macao	MAC	5660	454460
Macedonia	MKD	4794	598070
Madagascar	MDG	7880	164500
Malawi	MWI	7970	164540
Malaysia	MYS	5570	454580
Maldive Islands	MDV	5683	453600
Mali	MLI	7450	164660
Malta and Gozo	MLT	4730	574700
Marshall Islands	MHL	6810	368960
Martinique	MTQ	2839	353120
Mauritania	MRT	7410	164780
Mauritus	MUS	7850	164800
Mayotte	MYT	7881	164500
Mexico	MEX	2010	334840
Micronesia, Federated States of	FSM	6820	368960
Moldova, Republic of	MDA	4641	584980
Monaco	MCO	4272	532500
Mongolia	MNG	5740	484960
Monserrat	MSR	2485	356580
Morocco	MAR	7140	135040
Mozambique	MOZ	7870	165080
Namibia	NAM	7920	117100
Nauru	NRU	6862	722420
Nepal	NPL	5360	455240
Netherlands Antilles	ANT	2771	355320
Netherlands, The	NLD	4210	535280
New Caledonia	NCL	6412	725400
New Zealand	NZL	6141	715540
Nicaragua	NIC	2190	345580
Niger	NER	7510	165620
Nigeria	NGA	7530	165660
Niue	NIU	6144	715540
Norfolk Island	NFK	6022	710360
North Korea	PRK	5790	484080
Norway	NOR	4039	555780
Oman	OMN	5230	445120
Pakistan	PAK	5350	455860
Palua	PLW	6830	368960
Panama	PAN	2250	365900
Papua New Guinea	PNG	6040	725980
Paraguay	PRY	3530	336000
Peru	PER	3330	336040
Philippines	PHL	5650	456080
Pitcairn Island	PCN	6225	722960
Poland	POL	4550	586160
Portugal	PRT	4710	536200
Qatar	QAT	5180	446340
Reunion	REU	7904	166380
Romania	ROU	4850	586420
Russia	RUS	4621	586430
Rwanda	RWA	7690	166460
San Marino	SMR	4751	533800
Sao Tome and Principe	STP	7644	166240
Senegal	SEN	7440	166860
Serbia and Montenegro	SCG	4799	598910
Seychelles	SYC	7800	166900
Sierra Leone	SLE	7470	166940
Singapore	SGP	5590	457020
Slovakia	SVK	4359	587030
Slovenia	SVN	4792	597050
Soloman Islands	SLB	6223	722960
Somalia	SOM	7700	167060
South Afica, Republic of	ZAF	7910	117100
Spain	ESP	4700	537240
Sri Lanka	LKA	5420	451440
St. Helena	SHN	7580	166540
St. Kitts-Nevis	KNA	2483	356580
St. Lucia	LCA	2487	356580
St. Pierre & Miquelon	SPM	1610	226660
St. Vincent	VCT	2488	356580
Suadi Arabia	SAU	5170	446820
Sudan	SDN	7320	137360
Suriname	SUR	3150	367400
Svalbard & Jan Mayern Island	SJM	4031	555780
Swaziland	SWZ	7950	117100
Sweden	SWE	4010	557520
Switzerland	CHE	4419	557560
Syrain Arab Republic	SYR	5020	447600
Tajikistan	TJK	4642	467620
Tanzania	TZA	7830	168340
Thailand	THA	5490	457640
Togo	TGO	7520	167680
Tokelau Islands	TKL	6143	715540
Tonga	TON	6864	722420
Trinidad and Tobago	TTO	2740	357800
Tunisia	TUN	7230	137880
Turkey	TUR	4890	447920
Turkmenistan	TKM	4643	467950
Turks and Caicos Islands	TCA	2430	353880
Tuvalu Islands	TUV	6227	722960
Uganda	UGA	7780	168000
Ukraine	UKR	4623	588040
Unidentified	UNK	8220	999999
United Arab Emirates	ARE	5200	447840
United Kingdom	GBR	4120	538260
Uruguay	URY	3550	338580
USSR	USR	4610	688100
Uzbekistan	UZB	4644	468600
Vanuatu	VUT	6224	722960
Vatican Citty	VAT	4752	533800
Venezuela	VEN	3070	338620
Vietnam	VNM	5520	487040
Virign Islands (British)	VGB	2482	356580
Wallis and Futuna	WLF	6413	725400
West Bank (Israel)	PSE	5083	413760
Western Sahara	WSE	7370	135040
Western Somoa	WSM	6150	728882
Yemen (South)	YMD	5220	447200
Yemen Arab Republic	YEM	5210	448860
Yugoslavia	YUG	4790	598900
Zaire (Congo-Kinshasa)	COD	7660	161800
Zambia	ZMB	7940	168940
Zimbabwe	ZWE	7960	167160

Source: WISER.

Notes: U.S. column refers to the Census's schedule C for 2003. Schedule C is typically more detailed than United Nations country codes (column 4) but is subject to change over time.

U.S. DEPARTMENT OF COMMERCE – Economics and Statistics Administration – U.S. CENSUS BUREAU – BUREAU OF EXPORT ADMINISTRATION

FORM **7525-V** (7-18-2003)

SHIPPER'S EXPORT DECLARATION

OMB No. 0607-0152

1a. U.S. PRINCIPAL PARTY IN INTEREST (USPPI)(Complete name and address)

ZIP CODE

2. DATE OF EXPORTATION

3. TRANSPORTATION REFERENCE NO.

b. USPPI'S EIN (IRS) OR ID NO.

c. PARTIES TO TRANSACTION ☐ Related ☐ Non-related

4a. ULTIMATE CONSIGNEE *(Complete name and address)*

b. INTERMEDIATE CONSIGNEE *(Complete name and address)*

5a. FORWARDING AGENT *(Complete name and address)*

5b. FORWARDING AGENT'S EIN (IRS) NO.

6. POINT (STATE) OF ORIGIN OR FTZ NO.

7. COUNTRY OF ULTIMATE DESTINATION

8. LOADING PIER *(Vessel only)*

9. METHOD OF TRANSPORTATION *(Specify)*

14. CARRIER IDENTIFICATION CODE

15. SHIPMENT REFERENCE NO.

10. EXPORTING CARRIER

11. PORT OF EXPORT

16. ENTRY NUMBER

17. HAZARDOUS MATERIALS ☐ Yes ☐ No

12. PORT OF UNLOADING *(Vessel and air only)*

13. CONTAINERIZED *(Vessel only)* ☐ Yes ☐ No

18. IN BOND CODE

19. ROUTED EXPORT TRANSACTION ☐ Yes ☐ No

20. SCHEDULE B DESCRIPTION OF COMMODITIES *(Use columns 22–24)*

D/F or M (21)	SCHEDULE B NUMBER (22)	QUANTITY – SCHEDULE B UNIT(S) (23)	SHIPPING WEIGHT (Kilograms) (24)	VIN/PRODUCT NUMBER/ VEHICLE TITLE NUMBER (25)	VALUE (U.S. dollars, omit cents) (Selling price or cost if not sold) (26)

27. LICENSE NO./LICENSE EXCEPTION SYMBOL/AUTHORIZATION

28. ECCN *(When required)*

29. Duly authorized officer or employee

The USPPI authorizes the forwarder named above to act as forwarding agent for export control and customs purposes.

30. I certify that all statements made and all information contained herein are true and correct and that I have read and understand the instructions for preparation of this document, set forth in the "**Correct Way to Fill Out the Shipper's Export Declaration.**" I understand that civil and criminal penalties, including forfeiture and sale, may be imposed for making false or fraudulent statements herein, failing to provide the requested information or for violation of U.S. laws on exportation (13 U.S.C. Sec. 305; 22 U.S.C. Sec. 401; 18 U.S.C. Sec. 1001; 50 U.S.C. App. 2410).

Signature

Confidential – Shipper's Export Declarations (or any successor document) wherever located, shall be exempt from public disclosure unless the Secretary determines that such exemption would be contrary to the national interest (Title 13, Chapter 9, Section 301 (g)).

Title

Export shipments are subject to inspection by U.S. Customs Service and/or Office of Export Enforcement.

Date

31. AUTHENTICATION *(When required)*

Telephone No. (Include Area Code)

E-mail address

This form may be printed by private parties provided it conforms to the official form. For sale by the Superintendent of Documents, Government Printing Office, Washington, DC 20402, and local Customs District Directors. The "**Correct Way to Fill Out the Shipper's Export Declaration**" is available from the U.S. Census Bureau, Washington, DC 20233.

Figure 6. Shipper's export declaration (accessed April 12, 2006).

Table 15. Unrestricted regression by subsector by year

Description	NAICS	2001						1997					
		N	R^2	β_0	se	β_1	se	N	R^2	β_0	se	β_1	se
Food Prdcts	311	48	.83	0.84∗	0.31	0.86+	0.06	50	.83	0.41	0.32	0.90	0.06
Bev & Tbc	312	24	.71	1.91∗	0.31	0.60+	0.08	34	.56	0.73	0.43	0.76	0.12
Txt & Fbrcs	313	24	.39	2.21∗	0.73	0.56+	0.15	36	.55	2.45∗	0.29	0.47+	0.07
Txt Mill Prdcts	314	35	.62	0.24	0.41	0.91	0.12	42	.52	0.58	0.36	0.78	0.17
Apparel	315	22	.77	−1.71∗	0.52	1.33+	0.12	45	.69	−0.03	.40	0.93	0.10
Leather	316	43	.21	1.29	1.20	0.66	0.29	39	.58	0.74	0.37	0.76+	0.11
Wood Prdcts	321	41	.80	1.27∗	0.22	0.76+	0.06	49	.86	0.36	0.20	0.90	0.05
Paper	322	46	.76	0.80	0.41	0.87+	0.08	48	.86	0.11	0.28	0.96	0.06
Publishing	323	42	.72	−0.78	0.42	1.10	0.10	47	.48	1.50∗	0.35	0.72+	0.11
Petro & Coal	324	22	.63	0.43	0.64	0.84	0.15	41	.54	0.99∗	0.37	0.71+	0.10
Chemicals	325	43	.88	1.43∗	0.30	0.82+	0.05	50	.82	2.05∗	0.29	0.70+	0.05
Plscts & Rub	326	43	.76	1.21∗	0.36	0.80+	0.07	49	.85	1.21∗	0.23	0.77+	0.05
Nonmetal Mnrl	327	44	.77	1.37∗	0.26	0.73+	0.06	48	.85	0.48∗	0.22	0.89	0.05
Primary Metal	331	39	.73	0.36	0.50	0.92	0.09	48	.69	1.43∗	0.37	0.71+	0.07
Fab Metal	332	46	.77	0.67	0.38	0.85	0.07	50	.86	0.56∗	0.26	0.86+	0.05
Machinery	333	48	.88	0.84∗	0.32	0.89+	0.05	48	.90	0.16	0.31	0.97	0.05
Comps & Elect	334	42	.77	0.49	0.57	0.95	0.08	49	.87	0.66	0.34	0.91	0.05
Elect Eqpmnt†	335	42	.67	−0.35	0.65	1.04	0.11	45	.77	0.16	0.42	0.95	0.08
Trans Eqpmnt	336	46	.87	0.99∗	0.35	0.87+	0.05	50	.81	1.73∗	0.33	0.74+	0.05
Furniture	337	44	.81	0.96∗	0.19	0.75+	0.06	48	.65	0.52	0.27	0.78+	0.08
Misc Mnfcts	339	46	.82	0.20	0.36	0.97	0.07	49	.87	0.10	0.27	0.98	0.05

Sources: OM data from WISER; AR-1 data from *Exports From Manufacturing Establishments*, Census.

Notes: The model: $\log OM_{s,n,t} = \beta_{0_{n,t}} + \beta_{1_{n,t}} \log AR_{s,n,t} + \varepsilon_{s,n,t}$
(*) indicates significance at the 5% level. β_1 is significant for all NAICS; (+) indicates significantly different from 1 at the 5% level; Shapiro-Wilks rejects normality at the 5% level for 2001: 323 and 335 and for 1997: 312, 314, 323, 324, 325, and 337. Standard errors are not modified since robust standard errors do not change test conclusions; White's test does not reject null hypothesis (of homoskedasticity) for any subsector in 2001 but does for 326 and 337 in 1997. Non-fixed effects regression by state are not performed as the maximum number of observations is 21 and for small states, observations are extremely low. AK has one observation in 2001.

Table 16. Exports by location of port

State	2001		1997	
	Level	Percent	Level	Percent
CA	106.78	14.61	99.16	14.42
TX	95.00	12.99	76.18	11.08
NY	42.17	5.77	37.98	5.52
WA	34.93	4.79	32.75	4.76
MI	32.37	4.43	32.25	4.69
IL	30.43	4.16	26.45	3.85
FL	27.18	3.72	23.23	3.38
OH	27.09	3.71	24.90	3.62
NJ	18.95	2.59	15.17	2.21
MA	17.49	2.39	16.53	2.40
PA	17.43	2.38	16.07	2.34
NC	16.80	2.30	16.40	2.39
LA	16.59	2.27	18.73	2.72
GA	14.64	2.00	12.95	1.88
IN	14.37	1.97	12.03	1.75
US	731.03	100.00	687.60	100.00

Table 17. Exports by port

Port	2001		1997	
	Level	Percent	Level	Percent
Detroit, MI	56.93	7.79	51.21	7.45
JFK Airport	50.73	6.94	41.37	6.02
Laredo, TX	34.75	4.75	25.94	3.77
LAX Airport	34.73	4.75	37.35	5.43
SFI Airport	32.41	4.43	35.94	5.23
Buffalo, NY	31.08	4.25	38.83	5.65
New Orleans, LA	24.87	3.40	20.77	3.02
Houston, TX	21.25	2.91	22.28	3.24
Chicago, IL	20.97	2.87	19.07	2.77
Port Huron, MI	19.22	2.63	13.47	1.96
Long Beach, CA	17.75	2.43	12.22	1.78
New York, NY	16.35	2.24	7.42	1.08
El Paso, TX	16.13	2.21	10.10	1.47
Los Angeles, CA	15.76	2.16	24.21	3.52
Miami Airport	15.71	2.15	14.86	2.16
US	731.03	100.00	687.60	100.00

Source: WISER.

Notes: Levels are f.a.s. export sales of all commodities, not just manufacturers, in billions of current year dollar; U.S. total reflects exports from all "states" including Puerto Rico and unknown. Therefore totals may differ from previous tables. The eighteen state ports listed here account for 74.92% and 61.71% of exports in 2001 and 1997, respectively.

Table 18. Exports by foreign trade zone, 2001

State	Location	FTZ No.	Exports
TX	Austin	183	498.6
TX	McAllen	12	391.3
MS	Harrison County	92	323.1
FL	Miami	32	247.5
TX	El Paso	68	247.4
TX	Harris County	84	246.5
HI	Honolulu	09	217.0
GA	Atlanta	26	171.2
OK	Rogers County	53	147.7
FL	Broward County	25	129.6
TN	Nashville	78	116.8
MA	Boston	27	108.5
PA	Pittsburgh	33	104.0
PR	Mayaguez	07	93.9
TX	Laredo	94	85.8
US			3871.4

Source: 63rd Annual Report of the Foreign-Trade Zones Board to the Congress of the United States.

Notes: VA, given in millions of $2001, is based on material inputs, not sales or value-added; Each entry is a general FTZ located at or near a port of exit. FTZ sub-zones which are located at the site of production are not included. Sub-zones exported $11.6 billion in 2001, accounting for 75.32% of exports. Sub-zones are not included because since they are located at the site of production, exports through the sub-zone are attributed to the same state as the origin of production, and thus cause no problems for the OM series. Only exports through a general FTZ may cause problems since these are located at a port. Note a FTZ from Puerto Rico is included here, though PR is not included in determining the overall quality of the OM data.

Table 19. Manufacturing value-added by state, 2001 and 1997

State	2001		1997	
	VA	Percent	VA	Percent
AK	708	0.052	246	0.019
HI	813	0.061	858	0.052
WY	1224	0.091	889	0.069
MT	1249	0.093	1179	0.092
ND	1651	0.123	1657	0.129
SD	2753	0.205	2287	0.179
NV	2816	0.210	2296	0.179
VT	2958	0.221	2461	0.192
RI	3928	0.293	3946	0.308
DE	4049	0.302	3427	0.268
ME	4593	0.342	4566	0.357
NM	4717	0.352	9855	0.770
ID	4890	0.364	3739	0.292
WV	5155	0.384	6087	0.476
NH	6188	0.461	8133	0.635
US	1341330	100.000	1279823	100.000

Source: Bureau of Economic Analysis, `http://www.bea.gov/bea/regional/gsp/`.

Note: VA given in millions of current year dollars.