PRELIMINARY
COMMENTS WELCOME

Integrating Expenditures and Gross Product Originating Estimates: What To Do With the Statistical Discrepancy?*

J. Joseph Beaulieu (contact)
joe.beaulieu@frb.gov
Board of Governors of the Federal Reserve System
Industrial Output Section, Division of Research and Statistics
Washington, DC 20551

and

Eric J. Bartelsman
ebartelsman@econ.vu.nl
Free University of Amsterdam

March 2004

Abstract

The purpose of this paper is to investigate whether various disaggregated data can shed light on the possible sources of the statistical discrepancy. Our strategy is to use disaggregated data that sum to either GDP or GDI and compare the two in order to see where the discrepancy resides. To compare the two sets of data, we estimate two consistent sets of input-output models. We find a few "problem" industries that appear to explain most of the statistical discrepancy. A combination of the expenditure data and the income data, weighted more heavily towards the expenditure data, looks to produce the most sensible data according to a few economic criteria.

JEL Codes: C67, C82
Keywords: balancing, industry data, input-output, statistical discrepancy

* The paper was prepared for the CRIW Architecture for the National Accounts Conference, Washington, April 16-17, 2004. This paper represents the authors own views and not those of the Board of Governors of the Federal Reserve System or its staff.

A man with one watch knows what time it is;
A man with two watches is never quite sure.
—French Proverb

I. Introduction

The Bureau of Economic Analysis (BEA) publishes two measures of domestic output. The better known measure, gross domestic product (GDP), is the sum of private and government consumption and investment (including inventory investment) and net exports. A second measure, gross domestic income (GDI), is the sum of factor and nonfactor payments paid to input providers; these payments include compensation, profits and profit-like income, production and import taxes (formerly known as indirect business taxes), and the consumption of fixed capital. GDP and GDI conceptually measure the same thing, but because the two are calculated using imperfect source data, the two measures differ by what is called the statistical discrepancy.

Historically, the level of the statistical discrepancy has been small relative to GDP or GDI. As shown in the upper panel of chart 1, the absolute value of the statistical discrepancy as a fraction of the average of nominal GDP and nominal GDI peaked at 2.1 percent in 1993.

Nonetheless, different movements in real GDP and in real GDI can be economically meaningful. The bottom panel of chart 1 plots the average annual growth rates of real GDP and GDI. Although the movements of the two appear to coincide from year to year, between 1994 and 2000, real GDI grew on average ½ percentage point faster than real GDP, which is sizeable when compared to the average growth rate of the two series of 4.1 percent.

The recent difference in the growth rates of the two measures of domestic product is a problem for policymakers. The two measures imply different paths for productivity and potential output, which are important for planning purposes. Many analysts have pointed to the rapid rate of growth of GDI as being more consistent with the expected productivity gains from investment in high-tech equipment. Problems for analysts are especially acute when they need to combine data from the expenditure and income accounts, such as when modeling the components of national saving or projecting tax revenue. Indeed, the Congressional Budget Office points to the large swing in the statistical discrepancy as a substantial hindrance in its ability to forecast tax revenue in the past few years (CBO, 2003). The statistical discrepancy also leads to inconsistencies when analyzing particular types of income as a share of GDP.

Finally, the existence of the statistical discrepancy is a problem for researchers trying to reconcile their estimates of productivity trends by industry using data measured on the income side with aggregate estimates of productivity trends that are based on product-side measures. Bartelsman and Beaulieu (2004), Bosworth and Triplett (2003), and Nordhaus (2000) use the BEA's Gross Product by Industry data to model industry-level productivity. These data aggregate to GDI, making it hard to compare their results to the BLS's measure of productivity in the nonfarm business sector, which equals GDP less the value added from a few select sectors.

Several researchers have speculated on the data deficiencies that have led to the statistical discrepancy. GDP may be mismeasured because estimating the consumption of services is difficult (Council of Economic Advisers, 1997; Moulton, 2000) or exports are underreported (Moulton, 2000). GDI may be mismeasured because purging income of capital gains, which do not represent current production, is hard (Baker, 1998; Moulton, 2000), because stock options and other nontraditional forms of compensation show up in the compensation statistics without an offset in the profits data (Baker, 1998; Moulton, 2000), or because measures of proprietors' income have to be adjusted for underreporting in the tax return data. These adjustments to proprietors' income are based on an outdated and discontinued study (Council of Economic Advisers, 1997). Many of these explanations appear to be confirmed in Klein and Makino (2000), who find that the statistical discrepancy is inversely related to profits and proprietors' income and positively related to government spending and exports.[1]

The BEA prefers GDP as its measure of domestic output. Parker and Seskin (1997) write:

> [The BEA] considers the source data underlying the estimates of GDP to be more accurate. For example, most of the annual source data used for estimating GDP are based on complete enumerations, such as the Federal Government budget data, or are regularly adjusted to complete enumerations, such as the quinquennial economic censuses and census of governments….For GDI, only the annual tabulations of employment tax returns and Federal Government budget data are complete enumerations, and only farm proprietors' income and State and local government budget data are regularly adjusted to complete enumerations. For most of the remaining components of GDI, the annual source data are tabulations of samples of income tax returns.

This view is reflected in the presentation of the NIPAs. The BEA presents only GDP-related data in its summary tables, and in its decomposition of national income, it portrays the statistical discrepancy as if it were all an error in the measurement of income *vis-à-vis* GDP. A few years ago, the BLS appeared to adopt this view when it switched its definition of nonfarm business output in its Productivity and Cost release from one based on GDI to one based on GDP.

[1] Recall the convention that more GDP relative to GDI leads to a more positive statistical discrepancy; more GDI leads to a more negative discrepancy.

Others, however, have found GDI to have more desirable properties, at least at certain points in time. The Council of Economic Advisers (1997) found that the behavior of Okun's law, the rapid rise in personal tax payments, and the behavior of the real product wage was more consistent with the faster growing GDI measure of output in the mid 1990s, as measured at that time. During that same period, Greenspan (2004) observed that the rapid rise in measured labor and capital income, along with quiescent price inflation, suggested that productivity was increasing briskly. These productivity gains were apparent in the income-side measure, but not in the product-side measure of domestic output. Based on their time-series properties, Weale (1992) finds that GDI should be weighted almost twice as much as GDP in an optimal combination of the two measures into a single output series.

The main purpose of this paper is to conduct "forensic" exercises on the statistical discrepancy. This is done by confronting the various sources of disaggregated data in an integrated framework. Essentially, we allow the disaggregated data to sum to either GDP or GDI and compare the two in order to see where the discrepancy resides. In section 2 we present a detailed overview of the underlying source data and the method of integration. In section 3, we compare the differences in our industry estimates controlled to GDP data and estimates controlled to GDI data. It appears that the mismeasurement of deliveries to final demand and value added in a few problem industries explains most of the broad movements in the aggregate discrepancy. In the following section, we discuss the metrics used to find an optimal combination of the GDP and GDI data in one consistent set of industry estimates built on both GDP and GDI data. The fifth section concludes.

II. Method and Data

The basic strategy in this paper is to use disaggregated data that sum to either GDP or GDI and compare the two in order to see where the discrepancy resides. The exercise primarily involves reconciling nominal data; issues concerning how price and quantity indexes in the BEA's industry data, along with the statistical discrepancy, can be aggregated to obtain real GDP are considered in Moyer, Reinsdorf, and Yuskavage (2004). The two basic pieces of data that form our starting point are estimates of deliveries to final demand by industry and value added by industry. Deliveries to final demand are split among the major expenditure categories, consumption, investment, *etc.*, which sum to GDP. The value-added data aggregate to GDI.

Industry data on deliveries to final demand and value added are not directly comparable. Nothing in economic theory relates the two, except that the sum of each measure over industries should coincide. In order to have a basis for comparison, we estimate a full input-output system for each of the two sets of estimates. This allow us to compare the value added by industry implied by the published expenditure data to the published measure of value added by industry, and it allows us to compare the implications of the published value-added-by-industry data for deliveries to final demand to the published measure of deliveries to final demand. The input-output system that we estimate is described in more detail in section *II.1*.

In order to employ our empirical strategy, we need initial estimates of value added by industry, deliveries to final demand by industry, and gross output by industry. Not all of these data are published by government agencies, and so, some estimates had to be developed; other data needed to be reconciled to consistent concepts. Section *II.*2 describes the sources of these initial estimates.

To estimate a full input-output system controlled to each set of data, we convert the initial estimates to final estimates so that they are consistent with the linear restrictions embodied in the input-output theory. The conversion of the initial estimates to the final estimates is called balancing; the balancing routine that we employ is described in section *II.3*.

II.1. Input-Output System

The basic input-output system that we estimate is described in Figure 1. Domestic industries, represented as the first *N* rows of the table, produce gross output (vector *Y*) and deliver it to final demand (matrix *F*) or to other domestic industries, (columns of matrix *I*), who use it as intermediate inputs in their production processes. The fact that, for each industry, the sum of its deliveries to final demand and to other industries equals its gross output is called *the gross output identity*. The value added of an industry equals its gross output less the sum of its use of intermediate inputs (*value added identity*). The sum across industries of deliveries to final demand equals GDP (*GDP identity*), and the sum of value added across industries equals GDI (*GDI identity*). The *reconciliation identity* that integrates the system is that GDP equals GDI.

The first *N* rows of the system represent flows of goods, or expenditures on goods, from domestic industries. In order to simplify the exposition of our analysis, we account for the flows of imported goods in a nonstandard fashion: imported goods that are used in the production process of domestic industries or that are delivered to final domestic purchasers are part of a separate industry, called *Not Domestic Production*.

Deliveries of imports to domestic industries or to domestic purchasers are positive entries in the input-output system. The final demand category, imports, has an offsetting negative entry, so that the gross output of imports is zero. Note that, by definition, domestic industries do not deliver any output to the final demand category, imports, and so, the first *N* rows of the import column contain zeros.

Figure 1

Gross output identity: $I \bullet j_N + F \bullet j_5 = Y$ *where* $j_K =$ Kx1 vector of ones

Value added identity: $Y' \bullet Z - j'_{N+1} \bullet I = V$

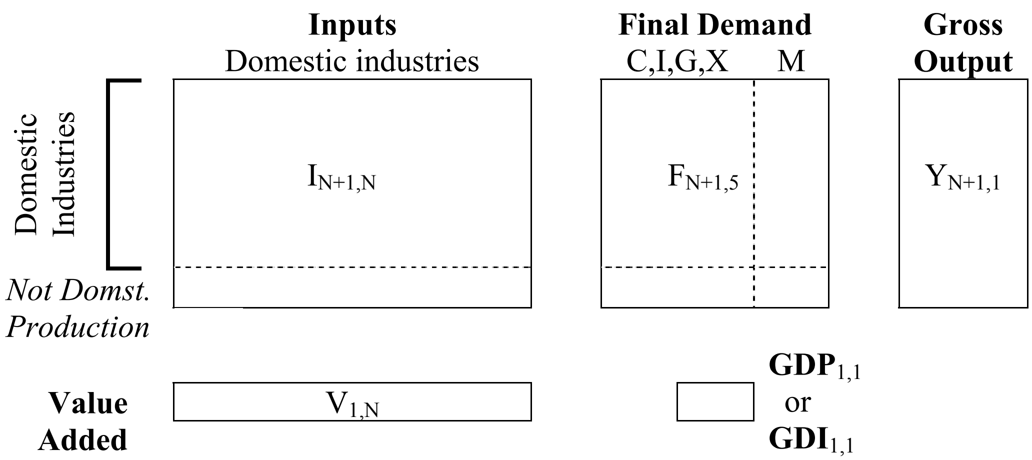
GDP identity: $j'_{N+1} \bullet F \bullet j_5 = GDP$

GDI identity: $V \bullet j_N = GDI$

Reconciliation: $GDP = GDI$

$$Z = \begin{array}{c|cccc} & 1 & 0 & \ldots & 0 \\ & 0 & 1 & \ldots & 0 \\ \text{N rows} & 0 & 0 & \ldots & 0 \\ & 0 & 0 & \ldots & 1 \\ \hline \text{1 row} & 0 & 0 & \ldots & 0 \end{array}$$

N+1 x N matrix



In addition, used and secondhand goods and scrap show up in the input-output accounts. They are used as intermediates to the production process and are either delivered to or supplied by the final demand categories. They do not represent new production, so like imports, their gross output equals zero. Negative entries represent net suppliers of the goods; positive entries represent net users. For example, businesses scrap some of their equipment each year, and so, the final expenditure category, business fixed investment, is a net supplier of used and secondhand goods and scrap. These commodity flows are also included in the pseudo industry *Not Domestic Production*.

I I.2. Initial data requirements

In order to develop our analysis, we need initial estimates of GDP, GDI, value added by industry (*V*), deliveries to final demand by industry (*F*), gross output by industry (*Y*), and intermediates used and supplied by industry (*I*). The GDP and GDI data for the years 1977 through 2001 come from the recently released benchmark NIPA data. Other data were adapted or created to be consistent with the latest NIPA aggregates.

II.2.a. Value added by industry

Value added for farms, private households, and owner-occupied housing, come directly from the NIPAs. Value added for owner-occupied housing was subtracted out of the real estate industry and placed in its own industry (before further aggregation). For other industries, estimates of value added by industry are sourced from the BEA's Gross Output Originating data set. Pre-1987 data were concorded to the 1987 SIC as in Bartelsman and Beaulieu (2004). All of the data were adjusted so that they sum to the latest estimates of gross domestic income.[2]

Value added in the real estate industry was also adjusted to exclude the imputed rental value of capital equipment and structures owned by nonprofit institutions. Instead, this imputed income was distributed to industries according to estimates of the compensation paid by nonprofit institutions by industry, as estimated in Bartelsman and Beaulieu (2004). Redistributing this income is useful because the final expenditures on many of the products produced by nonprofit institutions are not identified as to whether

[2] When the GPO data are updated in the summer, they will be used instead of these current estimates.

they were produced in the nonprofit sector or in the business sector, and so, these expenditures will not show up as coming from the real estate sector.

II.2.b. Deliveries to final demand by industry

For NIPA expenditure categories, except investment in software, structures, and inventories, the basic strategy to estimate deliveries to final demand by industry is first to map detailed expenditure categories in the NIPAs to the input-output tables' commodity classification system. These mappings are called "bridge tables", the construction of which is described in detail below. Bridge tables are first calculated for two years, 1987 and 1992 and then extended backwards and forwards in time assuming that the expenditures on each commodity grows with an estimate of its production.[3]

The delivery of each commodity to final demand is then estimated by allocating each NIPA expenditure category to the various commodities using the proportions in the constructed bridge tables. The domestic production of each commodity is then converted to an industry basis using the 1987 and 1992 make tables, and these industries are then aggregated to the definitions in Appendix A. Imports of all commodities are aggregated into one industry, called *Not Domestic Production*.

The method used to estimate the bridge tables differs by expenditure category. For personal consumption, detailed bridge tables were produced by the BEA for 1987 and 1992 that map 136 specific consumption categories to 307 I-O commodities. Two expenditure categories that show up as negative consumption, personal remittances in kind to nonresidents and expenditures in the U.S. by foreigners were excluded because instead of being mapped to industries, they were netted out of specific export categories.

---

[3] Estimates of the gross product of each commodity were created using gross output by industry at roughly a three-digit level mapped to commodities using the 1987 and 1992 make tables.

The available detail in the bridge table means that on average, each consumption category is divided among less than three commodities. On the other hand, commodities are produced by multiple industries. Chart 2a describes how well the PCE expenditure categories map to the 26 industries considered here. The top-left panel presents the number of expenditure categories whose commodities are produced by one industry, two industries, and so on. As indicated by the bars, the commodities of only 28 expenditure categories out of 138 are produced by one industry; 34 categories are composed of the product of 7 or more industries. The top-right panel makes a similar comparison; only this time, instead of counting the number of industries, it represents the fraction of PCE accounted for. As shown by the first bar, 30 percent of PCE is in expenditure categories whose products are produced by one industry, but roughly 32 percent is composed of categories whose products are produced by 7 or more industries.

The top panels, however, are not the whole story. Although most categories are composed of the product of multiple industries, typically most of an expenditure category is produced by a single industry. The bottom panels consider how much of an expenditure category is accounted for by a single industry. As shown in the bottom-left panel, for 82 categories, 95 percent or more of their product is produced by one industry. For another 13 categories, their dominant industry accounts for 90 to 95 percent of the expenditure category. The bottom-right panel translates these counts into shares of PCE. Almost 80 percent of PCE is accounted for by categories where 90 percent or more of its product is produced by one industry. All told, these results suggest that conditional on the expenditure categories being well measured, deliveries to personal consumption by industry should be estimated fairly precisely.

Bridge tables for investment in residential and nonresidential equipment (excluding software) are also published by the BEA. These tables map 27 expenditure categories to 118 commodities.[4] As shown in chart 2b, which is similar to chart 2a, multiple industries deliver to most of the expenditure categories, but for most industries, the vast majority is accounted for by a single industry.

The bridge table for the exports of goods was calculated differently. The 11 expenditure categories in the NIPAs were first split into 125 categories using data from the Census report on International Trade in Goods and Services that indicate how to apportion the NIPA data to the more detailed Census categories. Each I-O commodity was then assigned to one of the 125 Census categories. Expenditures of each of these categories were then allocated to each I-O commodity using the 1987 and 1992 use tables that indicate how much each commodity was exported. Margins for wholesale trade and goods transportation were assumed to be proportional to exports of all goods categories.

The bridge table for service exports was calculated by assigning commodities to one of the 7 expenditure categories and then allocating each NIPA expenditure category across assigned categories using data from the 1987 and 1992 use tables. Exports of travel and some of passenger fares and other services were adjusted to net out the two negative PCE categories.

The bridge tables for imports were calculated similarly to exports, but in this case, the 12 imported goods expenditure categories were split to 133 Census categories from the International Trade in Goods and Services. Imported services in the IO use tables were assigned to one of the seven imported service categories; imports of defense

---

[4] Tractors were pulled out of agricultural machinery and out of construction machinery using pre-revision NIPA data in order to conform to the BEA's published bridge table.

services and other services were assigned to noncomparable imports and imports of used goods and scrap.

Imports are different than other expenditure categories in that all imports are counted as coming from one industry. However, it is necessary to allocate a fraction of imports to the domestic final purchases categories and the rest to intermediate inputs to domestic production. This split was done by assuming that the fraction of an imported commodity delivered to various final demand categories versus to domestic industries is the same as that observed in the IO use tables that combine imported commodities and domestically produced commodities.

Bridge tables for government consumption were built by first assigning the consumption of fixed capital and the compensation paid to general government employees, excluding own-account investment to the general government industry. This category accounted for 62 percent of government consumption and investment in 1992. Compensation paid to employees for own-account investment is treated with other government investment. Commodities with positive value in the I-O use tables were assigned to government purchases of intermediate durables, nondurables, and services, depending on the commodity's characteristics. Commodities with negative values were assigned to government sales.[5] Netting out government sales from intermediate purchases yields government consumption excluding its own value added. The consumption of federal nondefense nondurable consumption was augmented with data from the Energy Information Agency to account for purchases and sales from the Strategic Petroleum Reserve.

---

[5] The NIPAs provide more detail on intermediate purchases for federal defense and the sales by state and local governments that are used to refine these assignments.

Bridge tables for government investment were created by first splitting own-account investment into equipment and structures using pre-revision data on compensation paid to force-account construction. Own-account investment originates from the general government. The remaining investment in structures was assigned to the construction industry, and the remaining investment in equipment was split among commodities using relative proportions in the 1987 and 1992 I-O use tables.

Residential and nonresidential investment in structures by industry had to be estimated in a different manner than would follow from the published input-output tables. Some expenditure categories were assigned directly to specific industries: drilling and exploration to mining, mobile homes to the appropriate manufacturing industry, and commissions to real estate.

The I-O tables appear to suggest that the remainder of investment in structures originates in the construction industry, but this is not correct. For construction, the I-O tables make an exception to the rule that production is classified according to the primary output of an establishment. Instead, the tables classify all construction regardless of the primary output of an establishment to the construction industry, a classification scheme known as activity based. Most of the rest of the input-output data are essentially organized on an establishment basis.[6] Chart 3 illustrates the problem with mixing establishment-based classifications and activity-based classifications: domestic investment in structures, excluding own-account investment exceeds the BEA's estimate of gross output in the construction industry. Consequently, we have to estimate how

[6] Farms and real estate services are the other industries in the input-output tables that are defined on an activity basis instead of an establishment basis. The farm industry is consistently treated in the NIPAs, however, and it is unlikely that much output counted in the farm sector is also counted elsewhere. All

much of private structures investment originates in the construction industry versus other industries.

The value of deliveries to final demand by the construction industry was calculated as a fraction of BEA's estimate of gross output. This equals the interpolated values of one minus the ratio of receipts for maintenance and repair to total sales in the Censuses of Construction (1977, 1982, 1987, 1992, and 1997).

The remainder of investment in structures was assigned to other industries based on their share of employment of construction workers in 2001 (from the BLS occupational survey) times BEA's estimate of the real wealth stock of structures by industry. Including the real wealth stock allows the indicators used to allocate the estimate of force-account construction to vary over time. Presumably, this estimate of force-account construction by industry is not included in the estimate of gross output, so the gross output of each industry was boosted by its estimate of construction activity.

Software investment was allocated across industries by first splitting investment into two components: own account and purchased software using BEA's detailed new investment-by-industry data. Own-account investment was then allocated across industries using these data. Purchased software was distributed to industries using the 1987 and 1992 make tables; 98 percent of the production of purchased software in 1992 was assigned to the data-processing services industry, SIC 737.

Inventory investment was allocated to industries based on published NIPA data. Farm inventories were assigned to farms. Manufacturing inventory investment was allocated across manufacturing industries using ASM book value data. ASM data

---

royalty income is counted in the real estate industry, but this the same treatment in the GPO data. Thus, it is unlikely that any adjustments would improve the consistency of the estimates.

reported on a NAICS basis or on the 1977 SIC were concorded to the 1987 SIC using available concordances. Wholesale and retail trade inventories were simply assigned to the trade industry. The remainder of inventory investment was assigned to other industries using data from the Sources of Income (Department of Treasury) for 1995-1997. Shares for other years were assumed to equal either the 1995 or 1997 value.

II.2.c. Gross output by industry

Estimates of gross output by industry come mainly from the GPO data, except for farms, owner-occupied housing, general government, and households, which are available or easily estimated from NIPA data. The gross output data were also adjusted for force-account construction, as described above.

In a few early years, the estimate of value added by the legal services industry was higher than the estimate of gross output. To allow our analysis to proceed, we boosted the value of gross output so that it exceeds value added by at least 5 percent, a figure consistent with the 1987 I-O use table.

II.2.d. Intermediate inputs

Unlike the vectors and matrices for gross output, deliveries to final demand, and value added, initial values for the intermediate block, *I*, are calculated only for the base years 1982, 1987, and 1992. Initial values for other years are developed iteratively using results from the balancing routine described below.

Initial values for the base year were calculated twice and then averaged. The first estimate allocates the vector of gross output less deliveries to final demand (*Y-F*) across the columns of *I* in proportion to the values observed in the 1982, 1987, or 1992 use tables. The second estimate allocates the vector of gross output less value added (*Y'-V*) across the rows of I, also in proportion to the values observed in the corresponding use tables. These two estimates, one of which can be thought of as consistent with the expenditure-side data, the other as consistent with the income-side data, are then combined by taking a geometric average of the two values cell by cell.

II.3. Balancing Routine

With initial values for gross output, deliveries to final demand, value added, and intermediate inputs, we estimate updated values such that the various constraints in the input-output system are satisfied with values "close" to the initial estimates. In particular, where a bar denotes initial estimates, we solve:

(1)
$$\min_{\{Y_t,F_t,V_t,I_t\}} \sum_{i=1}^{I}\frac{1}{\gamma_{it}^{Y}}\left(Y_{it}-\overline{Y_{it}}\right)^2+\sum_{i=1}^{I}\frac{1}{\gamma_{it}^{F}}\left(F_{it}-\overline{F_{it}}\right)^2+\sum_{i=1}^{I}\frac{1}{\gamma_{it}^{V}}\left(V_{it}-\overline{V_{it}}\right)^2$$
$$+\sum_{i=1}^{I}\sum_{j=1}^{I}\frac{1}{\gamma_{ijt}^{I}}\left(I_{ijt}-\overline{I_{ijt}}\right)^2$$
$$s.t. \qquad Y_{it}=F_{it}+\sum_{j=1}^{I}I_{ijt}$$
$$Y_{jt}=V_{jt}+\sum_{j=1}^{I}I_{ijt},$$
$$where \qquad \gamma_{st}^{X}=\sigma^{X}\left|\overline{X_{st}}\right| \;\; for\; X\in\{Y,F,V,I\}.$$

The values for σ^X – the tuning parameters – are set in advance. In the cases when $\sigma^X = 0$, the reciprocal becomes a Lagrangian multiplier, and the fact that $X_{it} = \overline{X_{it}}$ becomes another constraint in the minimization routine. In cases when $\overline{X_{it}} = 0,\ \sigma^X = 0$. This solution technique is a straight forward generalization of the least-squares technique first proposed by Stone, Champernowne, and Meade (1942).

Other solution techniques have been used for similar problems. In particular, a popular routine is the so-called RAS iterative solution. This iterative technique works best in cases when the sums of the rows and columns of a matrix to be estimated are fixed in advance and agree, when all of the initial values are positive, and when there are no "tuning" parameters, like σ^X. The technique can be adapted to handle negative initial values (Gunluk-Senesen and Bates, 1988) and could be further generalized to handle tuning parameters. In this case, the estimating procedure becomes another minimization routine; the kernel to be minimized instead of being quadratic becomes an entropy kernel.[7] Likewise, we could have chosen a log-quadratic kernel, $\frac{1}{\sigma^X}\log(X_{it}/\overline{X_{it}})^2$; the fact that $\gamma_{st}^X = \sigma^X \left|\overline{X_{st}}\right|$ in our routine can be thought of as an approximation to the log model. We chose a quadratic kernel because of computational speed.

As noted above, our estimation procedure is dynamic in that our initial estimates of $\overline{I_{st}}$ depend on the final results for other years when $t \neq$ 1982, 1987, 1992. We first estimate the system for 1982 and then move backwards in time to 1977 and forwards in time to 1986, using the final estimate of $I_{t\pm1}$ as a basis for $\overline{I_{st}}$. Specifically $\overline{I_{st}}$ is calculated by adapting $I_{t\pm1}$ for demand changes in the various columns by multiplying each cell of $I_{t\pm1}$ by the ratio of real gross output of column *j* in period *t* to real output of *j* in period $t \pm 1$. The matrix $I_{t\pm1}$ is also adapted for price changes in the various rows by multiplying each cell by the ratio of the gross output deflators from the BEA's GPO data

[7] See Schneider and Zenios (1990) credit a Russian mathematician Bregman for this result, although the proof that the first-order conditions for the minimization problem yield the RAS iterative solution is not

set for row *i* in period *t* to the output deflator in period $t \pm 1$. The same process is repeated starting in 1987 for the years 1983-1991 and starting in 1992 for the years 1988-2001. This produces two sets of estimates for 1983-1986 and for 1988-1991; these estimates are averaged to obtain one series for 1977-2001.

III. Results controlled to the GDP or GDI data

Equation (1) was first estimated under two sets of tuning parameters. The first set: $\left[\sigma^{Y}=0; \sigma^{F}=0; \sigma^{V}=1; \sigma^{I}=1\right]$ is consistent with the expenditure-side data and leads to estimates that add to GDP. We allow the income-side value added estimates to inform the industrial sector, but with $\sigma^{V}=\sigma^{I}=1$ the routine treats the estimates of value added symmetrically with the initial estimate of *I*. The second set of tuning parameters: $\left[\sigma^{Y}=0; \sigma^{F}=1; \sigma^{V}=0; \sigma^{I}=1\right]$ is consistent with the income-side data and lead to estimates that add to GDI. In both cases, $\sigma^{Y}_{MISC.}=\sigma^{V}_{MISC.}=0$ because the income and gross output of these industries are already integrated between the expenditure and income accounts. Early experiments with the estimation procedure gave estimates for the *Not Domestic Production* industry that tended to drift. With both negative and positive values for deliveries of this series tied down only to sum to zero, the estimates of this industry can be volatile. As a result, for $X \in \{V, F, I\}$, $\sigma^{X}_{NDP}=.00001$ if $\sigma^{X}_{NDP} \neq 0$.

Chart 4 plots the difference of the two estimates for each industry's deliveries to final demand in the left column of the panel and the difference of the two estimates for

---

difficult; see for example Gunluk-Senesen and Bates (1988). Bartelsman and Beaulieu (2003) review some each industry's value added in the right panel. Consistent with the definition of the overall discrepancy, the chart plots the difference in the measure calculated from the expenditure data less the measure calculated from the income side. The economy-wide statistical discrepancy is also plotted in all of the panels.

For most industries, the industry discrepancies are small relative to the overall discrepancy. Three industries, however, stand out: *Machinery and Instruments*, *Trade*, and *Finance and Insurance*, where the pattern of deliveries to final demand and value added appear to move with the total discrepancy. Indeed, as shown in chart 6, the difference in value added of the combination of these three "problem" industries, moved up in the early 1990s and dropped sharply subsequently, more so than the total discrepancy. The coincidence with the discrepancy in deliveries to final demand is not as sharp. The difference in deliveries to final demand of the problem industries remained flat in the first half of the 1990s, but like value added, the difference dropped sharply after 1996.

The fact that these three industries, *Machinery and Instruments*, *Trade*, and *Finance and Insurance*, show up as problem industries is not surprising. The *Machinery and Instruments* industry has evolved significantly over the last twenty-five years as productivity growth in high-tech industries has been substantial. Profit swings have been significant, and the adjustment of industry profits from a firm basis to an establishment basis is probably difficult. The semiconductor industry is particularly challenging as several firms have become "fables." These firms develop products but contract out their production to overseas fabrication plants. Morgan Stanley estimates that about 15 percent of the industry's worldwide revenue derived from products outsourced to

of the properties of the choice of minimizing kernel.

different firms (Edelstone, *et al,* 2003); much of this figure represents U.S. firm contracting with overseas foundries. Morgan Stanley expects this share to double by 2010.

The difficulties with the *Trade* industry likely relate to the accounting for margins on products sold. To the extent that these differences represent margins on domestic products, there is a corresponding offset in the difference between the two measures in the domestic industries producing the output. If this is the reason for the discrepancy in the trade sector, then it cannot be a source for the economy-wide discrepancy. On the other hand, if the differences arise from different margins on imported products, difficulties in tracing these products from imports to deliveries to domestic purchasers could be a source of the overall discrepancy.

Finance and Insurance is clearly an industry fraught with measurement difficulties. A good deal of banking services is not explicitly charged for. Banks offer services like “free checking” to its customers because it can make money by lending the balances customers leave in their accounts; customers choose to deposit their money in banks instead of lending it at higher rates to take advantage of the convenience of checking. The BEA has made substantial improvements to its estimation of imputed bank service charges in PCE and government consumption to account for these services (Fixler, Reinsdorf, and Smith, 2003); however, the division of these services between final demand and intermediate inputs to business is probably still imprecise. The accounting for insurance services is likewise difficult. The same issue of imputed intermediation services arises in insurance. Moreover, the true value of insurance services are not realized only when claims are paid; there is a continual flow of services.

Over the long run, the difference in premiums paid less claims paid equals the services provided. How to estimate these services over time is a thorny problem; the BEA has also improved its measures of deliveries to final demand of property-casualty insurance in the latest revision (Chen and Fixler, 2003). On the income side, adjusting for capital gains has to be more difficult in the *Finance and Insurance* industry than in any other.

Two other industries show some important differences that are not related to the overall discrepancy. Deliveries to final demand of *Chemicals, Refining, and Rubber and Plastics* estimated from the expenditure side rises sharply relative to estimated deliveries from the income side starting in 1999. In addition, the difference in estimated value added of the *Construction* industry shows a choppy pattern; smoothing through the pattern, the difference in the two estimates of value added is large relative to the level of the average of the two estimates.

Chart 7 plots the difference in the estimates of total deliveries to final demand by major expenditure categories. Almost all of the statistical discrepancy appears to reside in PCE. Except in the year 2001, the two sets of estimates of fixed investment essentially coincide. In particular between 1994 and 2000, a sizeable positive difference in deliveries to fixed investment by the *Construction* industry offsets a similarly sized negative difference in deliveries to fixed investment by the *Machinery and Instruments* industry.

IV. Optimal Combination of the Data

We next consider whether there is a combination of the data that provides an optimal result with respect to a metric based on 'desirable' economic properties. The economic properties that we consider concerns:

- the equalization of returns to capital;
- the orthogonality of total factor productivity shocks; and
- the stability of the intermediate block.

Our strategy is to estimate a series of input-output systems under different assumptions for the tuning parameters, $\sigma^{X.}$ For the input-output systems estimated under a particular set of tuning parameter, we calculate a statistic to evaluate the performance of the estimates with respect to each of the three economic properties.

IV.1. Equalization of returns to capital

The idea that returns to capital should be equalized across industries is straightforward. Simple arbitrage requires industries with below-average returns to sell their capital to industries with above-average returns to take advantage of the more profitable activity. Of course, if capital cannot be adjusted instantaneously because of adjustment costs, a putty-clay technology, or the quasi-fixity of capital, then the simple arbitrage argument breaks down. The fact that we do not estimate equalized capital returns under any calculation suggests that something more than data mismeasurement is needed to explain cross-sectional variation in capital returns. Nonetheless, data mismeasurement probably widens the distribution of returns; estimates that minimize the variation are indicative of an optimal combination of the expenditure-side and income-side data with respect to this metric.

To measure the performance of each I-O estimate, we calculate the return to capital for each year for the industries that we consider, except *Government Enterprises, Miscellaneous Industries,* and *Not Domestic Production*. These industries are excluded because there is no presumption of profit maximizing behavior of government enterprises, the general government, or owner-occupied housing. For each year we calculate the variance of returns across industries and then average the variance over the 1977-2001 period.

The return to capital is defined as capital income divided by an estimate of the wealth stock. Capital income equals value added less compensation paid to all types of labor less non-capital taxes on production and imports plus government subsidies. These data come from Bartelsman-Beaulieu (2004) as adapted from the Gross Product Originating data. Compensation is adjusted to include an imputation for the labor income of the self employed. As measured in the NIPAs this income is counted in proprietors' income.[8] Non-capital production taxes are composed mostly of sales taxes. Simply plugging in the these data on compensation, taxes, and subsidies assumes that these components of income paid are not mismeasured. The compensation data, at least to employees, is probably better measured than profits, interest, and proprietors' income; nonetheless, the idea that all of the mismeasurement of income resides in capital income is simply a maintained hypothesis that is not pursued further.

Estimates of the wealth stock are calculated based on detailed BEA estimates of investment by industry and by asset type. Wealth stocks were calculated using the appropriate formula (Hulten, 1990) that is consistent with the age-efficiency schedule used in Bartelsman-Beaulieu (2004). The BEA investment data are adjusted for each input-output estimate of total investment to the extent that estimated deliveries to the final demand category private fixed investment differs from the original estimate in the NIPAs on which the detailed BEA data are based.

IV.2. Orthogonality of innovations to total factor productivity

The idea that variation in GDP is driven by productivity shocks that are common across industries is a central tenant of real business cycle theories. Opponents to this theory have generally held that the size of the aggregate shock required to generate business cycle variation is implausibly large; candidate sources for such aggregate shocks, such as the weather, appear to amount to little. Simply adding up idiosyncratic shocks leads to an aggregate productivity shock that does not equal exactly zero, but because of the law of large numbers the aggregate is too small unless the sector-specific shocks are large.[9]

Inherent in the counter argument to real business cycle models is that industry TFP growth rates should be uncorrelated. With measurement error, however, TFP growth rates can be correlated, even if they are orthogonal in reality. The measurement error can be correlated if it involves an allocation error of a fixed aggregate across industries. If the measurement error affects industries differently and this is somehow related to the business cycle — perhaps due to whether the product is a good or service — mismeasurement can also generate a correlation.

Economists have tested whether there is a common factor to industry shocks (Lebow, 1990; Forini and Reichlin, 1998). In this exercise we do the opposite: we

---

[8] The BLS makes the same adjustment in its Productivity and Cost estimates.

[9] Horvath (2000) shows that the law of large numbers has to be augmented by the input-output structure of the economy. If the input-output table is sparse, then the law of large numbers applies at a much slower rate than is commonly presumed.

assume that this common factor is small and look for what combination of data produces a set of TFP growth rates that are as close to orthogonal as possible. To measure the orthogonality of TFP growth rates, we calculate the contemporaneous correlation matrix for each set of input-output estimates and then collapse the correlation matrix into a single statistic by taking its determinant. The determinant is minimized when the correlation matrix equals the identity matrix and is larger the more correlated the TFP innovations are.

Industry TFP measures are calculated by modeling real gross output as a function of capital services, labor hours, and real intermediate inputs, using the usual Divisia formulation. Deflators for gross output come from the BEA's GPO data set, as adopted in Bartelsman and Beaulieu (2004). The same gross output deflators are used to generate a deflator for intermediate input usage. Industry data on hours and capital services also come from Bartelsman and Beaulieu, although capital services built from investment flows are adjusted for differences in estimated aggregate deliveries to business fixed investment, as in subsection *IV.1.*

IV.3. Stability of intermediate block

The idea that the coefficients of an input-output table should be stable is common in the literature. After all, the coefficients represent the structure and technology of an economy that evolve slowly due to "technical progress, exhaustion of natural resources, or variation in consumers' tastes"; the stability of the structure of the economy stands in contrast to final demand, which is less stable (Leontief, 1953). Immediately, the question arises whether the stability of input-output coefficients should be measured using nominal data or real estimates (see Sawyer, 1992 and references therein), and whether the values in the intermediate block should be constant with respect to the gross output of the supplying industries or the gross output of the demanding industries. De Mesnard (2002) uses the relative stability of the cells of the intermediate block divided by supplying industries versus those divided by demanding industries as a measure of whether an industry is "supply oriented" or demand oriented."

For each estimate of the input-output system, we make four different calculations: two use nominal data; two use real data. Real data are calculated by dividing the rows of the input-output table by gross output deflators from Bartelsman and Beaulieu (2004). When using deflated measures, we ignore the obvious complications of taking ratios of chain-aggregated deflated data (Whelan, 2002). Let $D(Y)$ denote a square matrix with the gross output vector *Y* along the main diagonal and zeros otherwise. Define allocation and technical coefficients as

$$\begin{aligned} &\text{Allocation coefficients:} \quad D(Y)^{-1} \cdot I \\ &\text{Technical coefficients:} \quad I \cdot D(Y)^{-1}. \end{aligned}$$

We then take the standard deviation across time of each cell and then collapse this matrix into a single statistic by taking a weighted average of the standard deviations of each cell, where the weights equal the average of the absolute value of the cells of *I* over time.

IV.4. Results

Chart 8 plots the results of these exercises. On the bottom axis of each panel are the values for $\left\{\sigma^F / \sigma^V\right\}$; in all cases $\sigma^Y = 0, \sigma^I = 1$. Two other I-O systems were calculated, denoted as $\{0/\infty\}$ and $\{\infty/0\}$. The first system, $\{0/\infty\}$, is calculated by sweeping the vector *Y-F* across the columns of the initial estimates of *I* without any reference to the initial values of *V*; the value for *V* is calculated as a residual according to the value added identity. The second system, $\{\infty/0\}$, is calculated by sweeping the vector *Y'-V* across the rows of the initial estimates of *I*, ignoring the initial values of *F*; the resulting value for *F* is calculated using the gross output identity.

The upper-left panel of chart 8 plots the average cross-sectional standard deviation of the return to capital. Except for the estimate {0/.5}, variation in the return to capital is similar across estimates; the variation for {0/.5} is almost 15 percent lower than the others.

The upper-right panel plots the determinant of the correlation matrix of the growth rates in TFP. The panel indicates that systems calculated using {0/1} and {0/.5} produce estimates that are fairly uncorrelated in comparison with the other estimates.

The upper-two panels involve calculations that use value added, total intermediate usage, and the like; they are consistent in pointing to the systems that emphasize the expenditure-side initial estimates of final deliveries, $\sigma^F = 0$, but also allow for an important role for the income-side value added initial estimates, $\sigma^V = .5$ or 1.

In contrast, the bottom panel emphasizes the individual cells of *I*, and they point to the opposite conclusion. The estimates for {0/1} and {0/.5} produce the greatest variation in allocation and technical coefficients. On the other hand, estimates that mix the expenditure-side and income-side data without being constrained to either appear to produce fairly stable coefficients. The fact that the $\{0/\infty\}$ and $\{\infty/0\}$ estimates produce the least variation in the standard deviation of real technical coefficients is essentially by construction because the calculation of the initial values of *I* are developed under the assumption that the real technical coefficients are constant.

V. Conclusion

In this paper we have employed industry estimates of deliveries to final demand and value added to investigate possible sources of the statistical discrepancy. We find that the expenditure-side data and the income-side data imply two different paths for the production of goods and services from the *Machinery and Instruments*, *Trade*, and *Finance and Insurance* industries that appear to be related to the statistical discrepancy. Important for the measurement of recent movements in productivity, there is an anomalous shortfall in 2001 in the change in deliveries to final demand of the *Machinery and Instruments* industry implied from the income-side data relative to the change in deliveries that is measured from the expenditure-side data. At a minimum, it might be useful to push on the source data for these industries to see if some improvement in data collection could help reconcile these discrepancies.

Our analysis also uncovered some other possible discrepancies that warrant some attention, even if they are not consistently related to the aggregate discrepancy. There are some important differences in our two sets of estimates of deliveries to final demand in the *Chemicals, Refining, and Rubber and Plastics* industry and in the *Communications* industry. There are also some differences in the estimates of value added in the *Construction* and *Real Estate* industries; these may be related and probably owe to some problems in allocating income paid between the two industries.

Across expenditure categories, we found that the statistical discrepancy may reside wholly in PCE. This may be due to the way sales data are collected and allocated among various expenditure categories versus supplied to other businesses. It may also be somewhat accidental. The discrepancy in fixed investment was small because of offsetting differences in deliveries to final demand by the *Machinery and Instruments* industry — whose discrepancy became significantly negative in the last half of the 1990s — and deliveries to final demand by the *Construction* industry — whose discrepancy became positive at the same time. This offsetting pattern is probably more coincidental than causal.

As a byproduct of this analysis we produced a consistent set of estimates of industry gross output, deliveries to final demand, intermediates used, and value added. Indeed, we produced a lot of consistent estimates and offered some means to judge how they should be combined. Some combination of the expenditure-side and income-side data should be employed, perhaps weighted more to the GDP data than the GDI data. Such a weighting scheme seems to provide the best mix over various criteria. Our methodology could easily be augmented if there was some *a priori* idea as to how the quality of the source data varies across sources (see Lawson *et al.*, 2004)

We could not have written this paper if the BEA had not produced the wealth and the variety of the data that it does. Besides all of the information provided in the NIPAs, the GPO data, and the published input-output tables, including the bridge tables, the importance of various estimates that the BEA makes available on its website for researchers, such as the tables on underlying expenditure detail and the estimates of investment by industry and by asset type should not be overlooked. Of course, there would be no point in writing this paper if the BEA did not publish two estimates of domestic product; some countries only produce one estimate by balancing the information from expenditure-side and income-side data. If the BEA published only one estimate of domestic product then only the BEA could have done the forensic analysis in this paper.

Even though "the man with two watches is never quite sure" what time it is, the man with one watch may not realize that his watch has slowed or even stopped. An English version of this proverb that we have seen starts with "It's possible to own too much …"; as economists we know this cannot be true, especially where data are concerned. Policymakers found important clues in the income-side measures of the transition of the economy when the production of and investment in high-tech goods pushed the growth rate of potential GDP higher (Jorgenson and Stiroh, 2000).

As part of its strategic plan, the BEA plans to publish integrated value-added I-O accounts with GDP-by-Industry accounts (U.S. Bureau of Economic Activity, 2003). These integrated data add to GDP (Lawson *et al.*, 2004); they will supplant the current Gross Output Originating data that add to GDI. While all of the research community would welcome a consistent, integrated data set that relates deliveries, inputs used, and measures of income paid, it needs to be recognized that separate industry data that are derived and are consistent with published estimates of GDI are valuable as a check against the integrated estimates, as well as against the aggregate GDP and GDI measures. After applying the appropriate concordance, researchers could also use these GDI-based data to extend the existing GPO data set.

It is easy to recommend that others find resources in their budgets to provide additional data. Fortunately, the BEA already publishes a lot of the data that would be needed to develop a set of industry estimates of value added that add to GDI. In Section 6 of the NIPAs – Income and Employment by Industry – the BEA provides data on income paid by industry. The problem with using these data directly is that some of the data are organized on a firm basis, instead of an establishment basis.[10] However, if the BEA were to make available on its website the factors that it uses to convert the data on a firm basis to an establishment basis — something the BEA will have to develop in-house anyway in order to prepare its integrated accounts — the research community could develop a second consistent data set in real time that could be used to monitor and investigate future data discrepancies.

[10] Specifically, corporate profits, nonfarm proprietors' income, net interest paid, and capital consumption allowances are published on a firm basis; see U.S. Bureau of Economic Analysis (2001), pages M21-M22. Bartelsman and Beaulieu (2004) describe the magnitudes of the differences between firm-based estimates and data converted to an establishment basis.

References

Baker, Dean (1998), "The New Economy Does Not Lurk in the Statistical Discrepancy", *Challenge* 41(4) (July/August), 5-13.

Bartelsman, Eric J. and J. Joseph Beaulieu (2004), "A Consistent Accounting of U.S. Productivity Growth", Mimeo, Board of Governors of the Federal Reserve System (April).

Bartelsman, Eric J. and J. Joseph Beaulieu (2003), "Techniques to Reconcile Data with Linear Growth", Mimeo, Board of Governors of the Federal Reserve System (January).

Bosworth, Barry P. and Jack E. Triplett (2003), "Services Productivity in the United States: Griliches' Services Volume Revisited, Mimeo, The Brookings Institution, (September), http://www.brookings.org/views/papers/bosworth/20030919.htm.

Chen, Baoline and Dennis J. Fixler (2003), "Measuring the Services of Property-Casualty Insurance in the NIPAs: Changes in concepts and methods", *Survey of Current Business* 83 (10) (October) 10-25.

Congressional Budget Office (2003), "CBO's Economic Forecasting Record: An evaluation of the economic forecasts CBO made from January 1976 through January 2001." (October), http://www.cbo.gov/showdoc.cfm?index=4639&sequence=0.

Council of Economic Advisers (1997), *Economic Report of the President*, Washington, D.C.: United States Government Printing Office.

Edelstone, Marc (2003), "Transition to 300-mm Wafers Should Drive Secular Changes", Equity Research, Morgan Stanley (December 1).

Fixler, Dennis J., Marshall B. Reinsdorf, and George M. Smith (2003), "Measuring the Services of Commercial Banks in the NIPAs: Changes in concepts and methods", *Survey of Current Business* 83 (9) (September) 33-44.

Greenspan, Alan (2004), "Risk and Uncertainty in Monetary Policy: Remarks at the Meetings of the American Economic Association", San Diego, CA (January 3).

Gunluk-Senesen, G. and J.M. Bates (1988), "Some Experiments with Methods of Adjusting Unbalanced Data Matrices", *Journal of the Royal Statistical Society Series A (Statistics in Society)* 151(3), 473-490.

Klein, L.R. and J. Makino (2000), "Economic Interpretations of the Statistical Discrepancy", *Journal of Economic and Social Measurement* 26(1) 11-29.

Lawson, Ann, Brian Moyer, Sumiye Okubo, Mark Planting (2004), "Integrating Industry and National Economic Accounts: First Steps and Future Improvements", Mimeo, Bureau of Economic Analysis.

Moulton, Brent R. (2000), "Getting the 21st-Century Right: What's Underway?", *The American Economic Review: Papers and Proceedings* 90(2) (May), 253-258.

Moyer, Brian C., Marshall B. Reinsdorf, and Robert E. Yuskavage (2004), "Aggregation Issues in Integrating and Accelerating BEA's Accounts: Improved Methods for Calculating GDP by Industry", *Mimeo*: Bureau of Economic Analysis (February).

Nordhaus, William D. (2000), "New Data and Output Concepts for Understanding Productivity Trends, Cowles Foundation Discussion Paper No. 1286, Yale University (November). http://cowles.econ.yale.edu/P/ab/a12/a1286.htm

Parker, Robert P. and Eugene P. Seskin (1997), "The Statistical Discrepancy", *The Survey of Current Business* 77(8) (August), p. 19.

Schneider, M.H. and S.A. Zenios (1990), "A Comparative Study of Algorithms for Matrix Balancing", *Operations Research* 38, 439-455.

Stone, Richard, D.G. Champernowne, and J.E. Meade (1942), "The Precision of National Income Estimates", *The Review of Economic Studies* 9(2) (Summer), 111-125.

U.S. Bureau of Economic Activity, Department of Commerce (2003). "BEA Strategic Plan for FY 2003-FY 2007", Mimeo, Bureau of Economic Activity, http://www.bea.doc.gov/bea/about/0307finalstratplan.pdf

U.S. Bureau of Economic Activity, Department of Commerce (2001). "A Guide to the NIPAs", Mimeo, Bureau of Economic Activity, http://www.bea.doc.gov/bea/an/nipaguid.htm

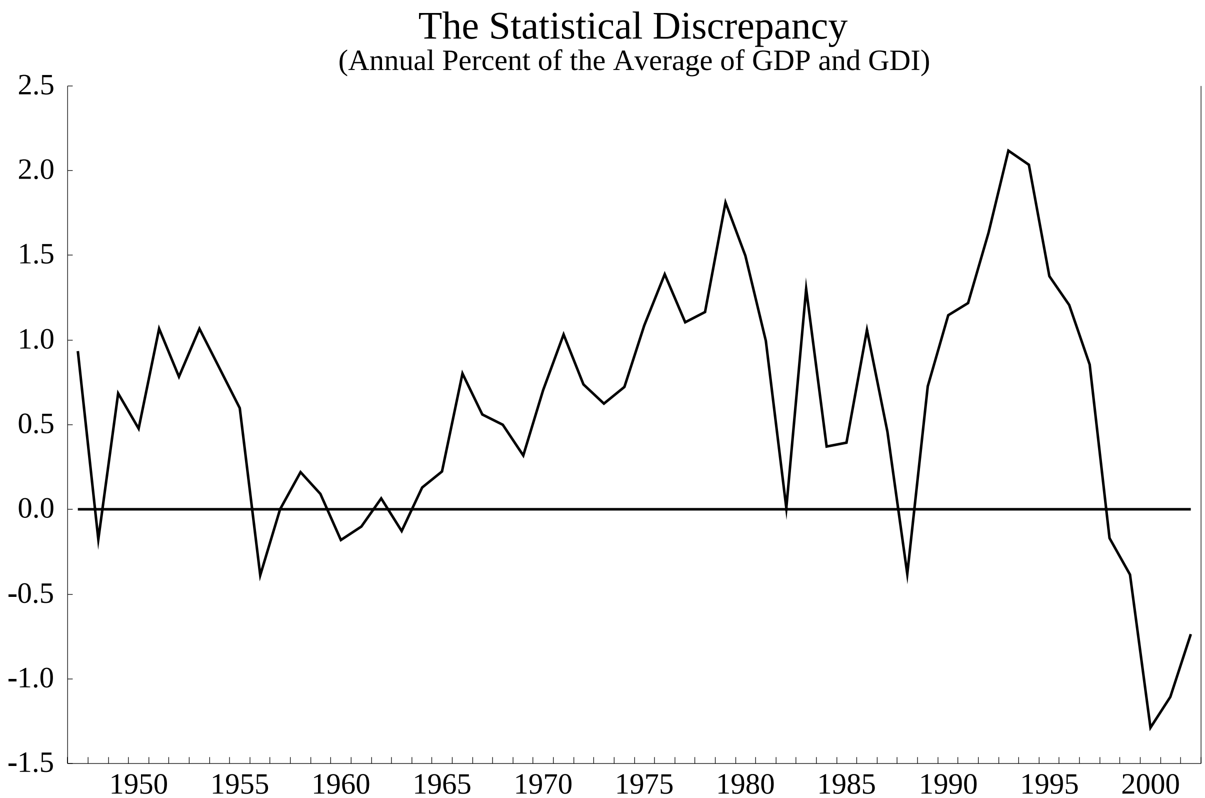
Weale, Martin (1992), "Estimation of Data Measured with Error and Subject to Linear Restrictions", *Journal of Applied Econometrics* 7(2) (April-June), 167-174.

Appendix A

Industry	SIC 87	Description
Agriculture	01-09	Farms, agricultural services, forestry, fishing, hunting, and trapping.
Mining	10-14	Metal mining, coal mining, oil & gas extraction, and mineral mining.
Construction	15-17	Construction.
Wood, Furniture, Paper, and Printing	24-27	Manufacturers of lumber and wood, furniture, paper, and printing.
Primary Durable Mfg.	32-34	Stone, clay and glass, primary metal and fabricated metal manufacturing.
Machinery and Instruments	35-36, 38-39	Machinery, electrical machinery, instruments, and miscellaneous manufacturing. The aggregate includes, computers, communications equipment, and semiconductors.
Transportation Equipment	37	Motor vehicles and parts, aircraft and parts, and other transportation equipment.
Food and Tobacco	20-21	Food and beverages and tobacco manufacturing.
Textiles, Apparel, and Leather	22-23, 31	Textiles, apparel, and leather manufacturing.
Chemicals, Refining, and Rubber & Plastics	28-30	Chemicals, petroleum refining, and rubber & plastics manufacturing.
Transportation	40-42, 44-47	Trucking, water, rail, and air transport, warehousing, pipelines (ex. natural gas), and transportation services.
Communications	48	Telephone and telegraph, radio and television, and other communications services.
Utilities	49pt.	Electrical, natural gas, and water and sanitary services utilities. It excludes government enterprises such as TVA and Bonneville.
Trade	50-59	Wholesale and retail trade.
Finance and Insurance	60-64, 67	Depository and nondepository institutions, securities dealers and brokers, insurance carriers and agents, and holding companies.
Real Estate	65	Real estate, excluding imputations for owner-occupied housing and the rental value of nonprofits' capital.

Industry	SIC 87	Description
Hotels and Other Lodging	70	Hotels and other lodging.
Personal Services	72, 75-76	Personal services, automotive repair services, and parking, and miscellaneous repair services.
Business Services	73	Business services, including software and data processing.
Movies and Recreation Services	78-79	Motion pictures, and amusement & recreation services.
Health Services	80	Health services.
Legal Services	81	Legal services.
Other Services	82-84, 86-87, 89	Social services, museums, membership organizations, engineering, accounting, research, and management services, and miscellaneous services.
Government Enterprises	43, 49pt, other	Federal and State and local government enterprises, including the Postal Service, TVA, and Bonneville Power.
Miscellaneous Industries	88, other	Private households, owner-occupied housing, and general government.
Not Domestic Production	—	Imports, used and secondhand goods, and scrap.

Chart 1


The Statistical Discrepancy
(Annual Percent of the Average of GDP and GDI)
2.5
2.0
1.5
1.0
0.5
0.0
-0.5
-1.0
-1.5
1950
1955
1960
1965
1970
1975
1980
1985
1990
1995
2000

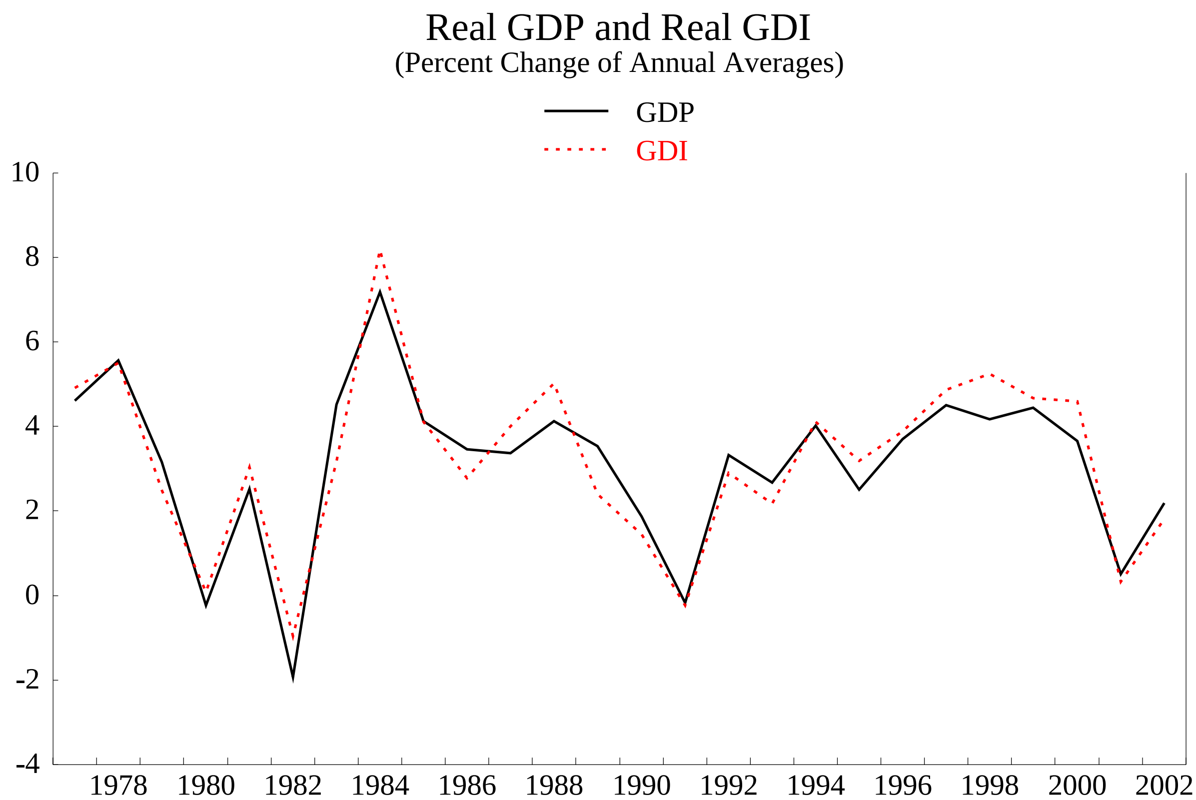

Real GDP and Real GDI
(Percent Change of Annual Averages)
GDP
GDI
10
8
6
4
2
0
-2
-4
1978
1980
1982
1984
1986
1988
1990
1992
1994
1996
1998
2000
2002

Chart 2a
Distributions of Mappings of Industries to NIPA Expenditure Categories
PCE

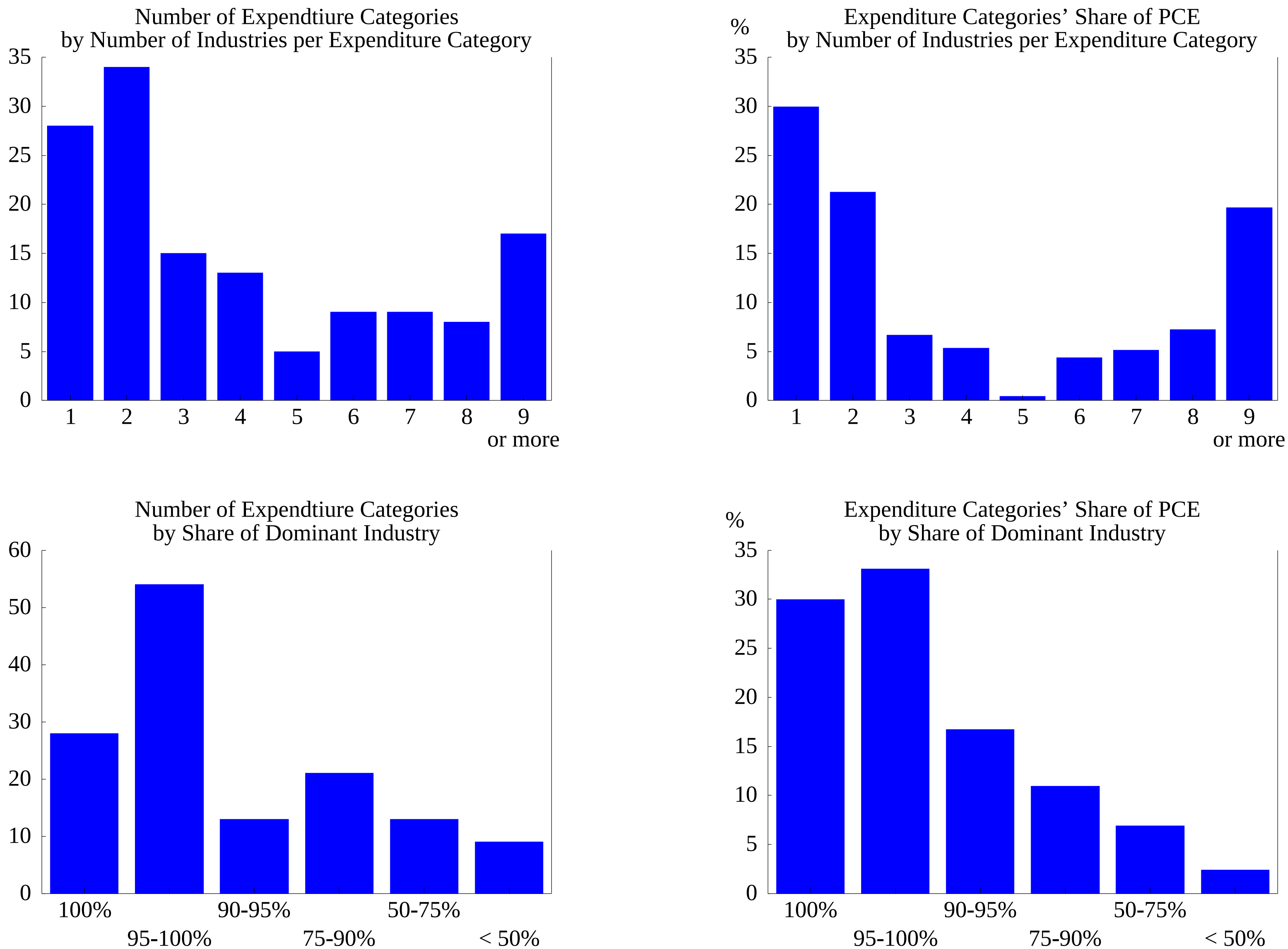


Chart 2b
Distributions of Mappings of Industries to NIPA Expenditure Categories
Equipment Investment

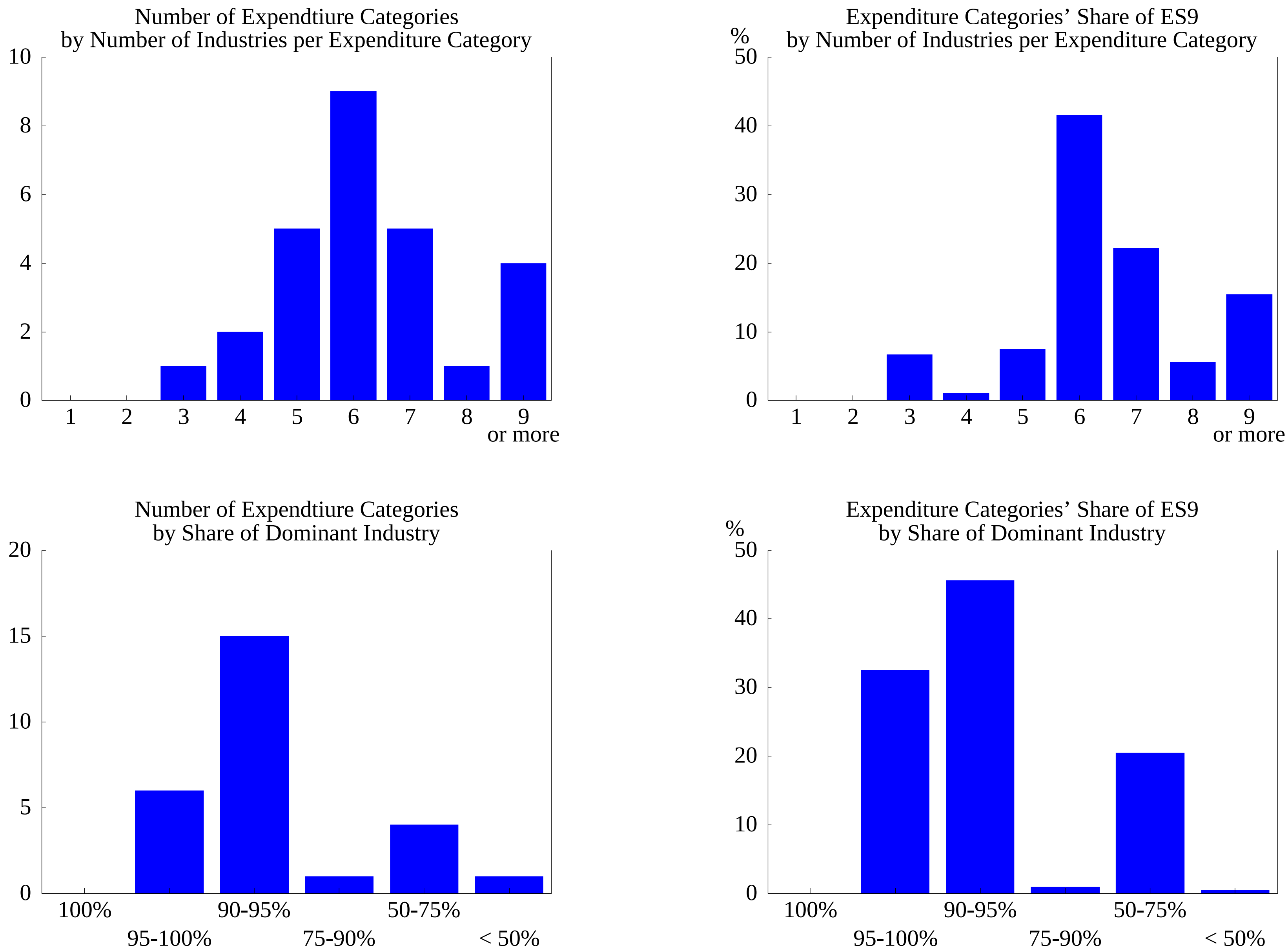
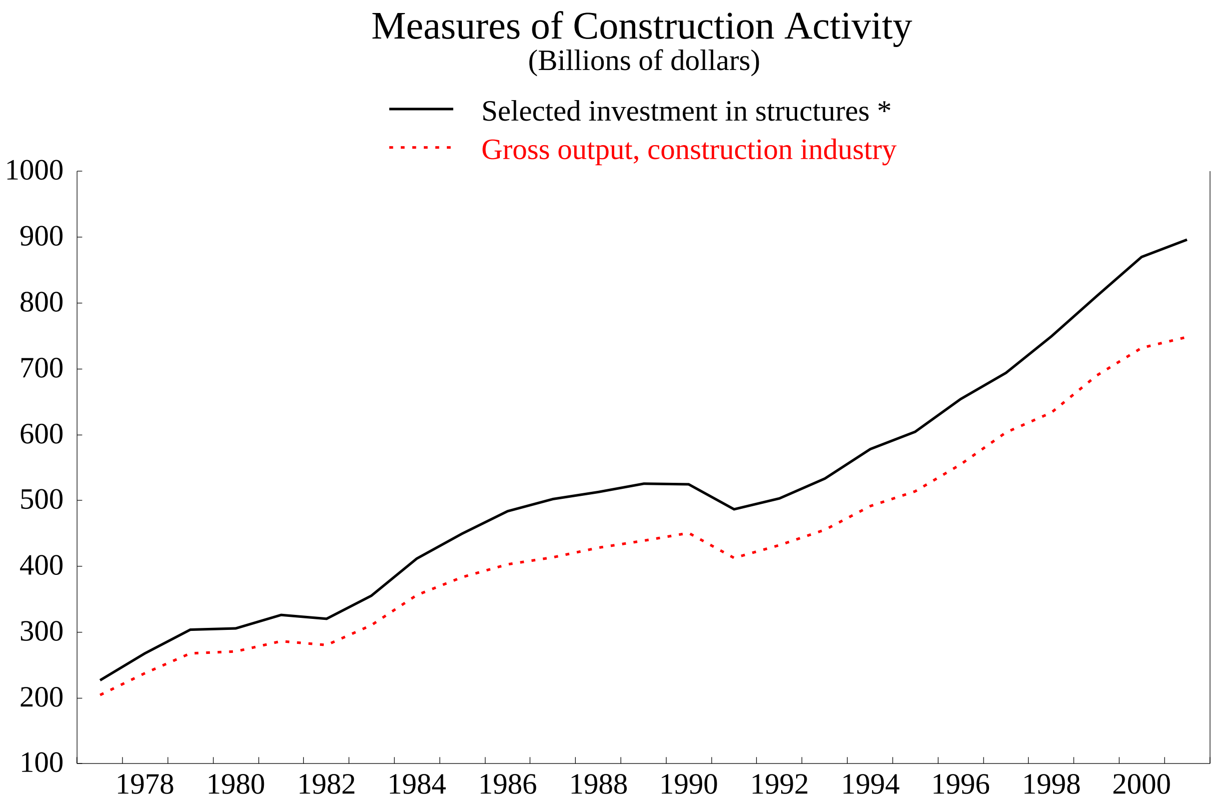


Chart 3



* Excludes investment in mining and exploration, manufactured homes, commissions, and government own-acocunt investment.

Chart 4
Statistical Discrepancy by Industry
(Billions of dollars)

Industry Discrepancy in Black (Thick)   Total Discrepancy in Red (Thin)

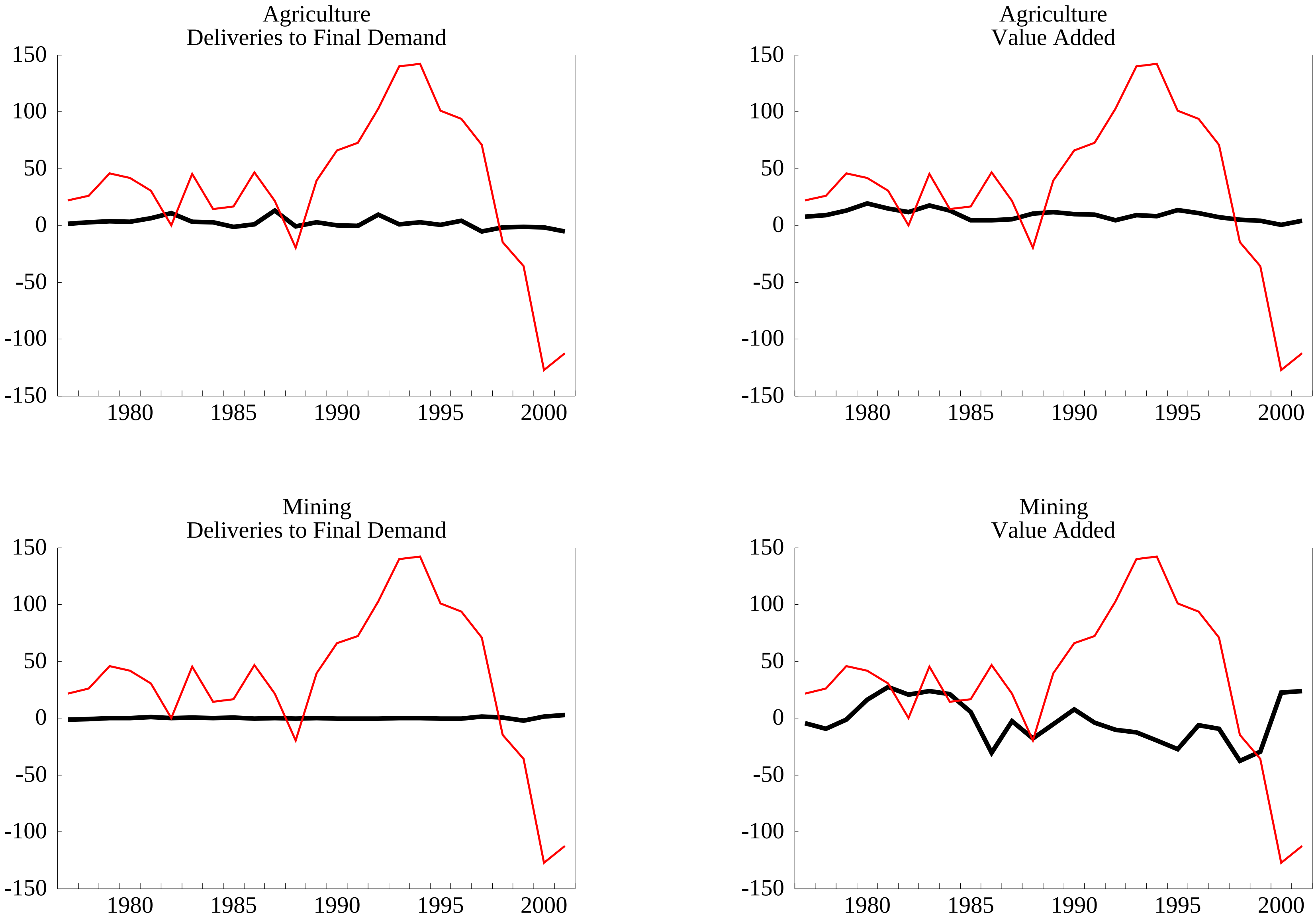
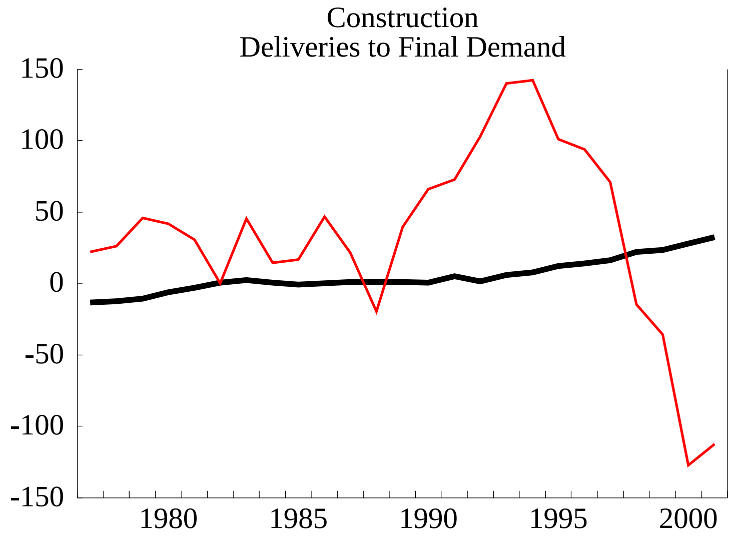
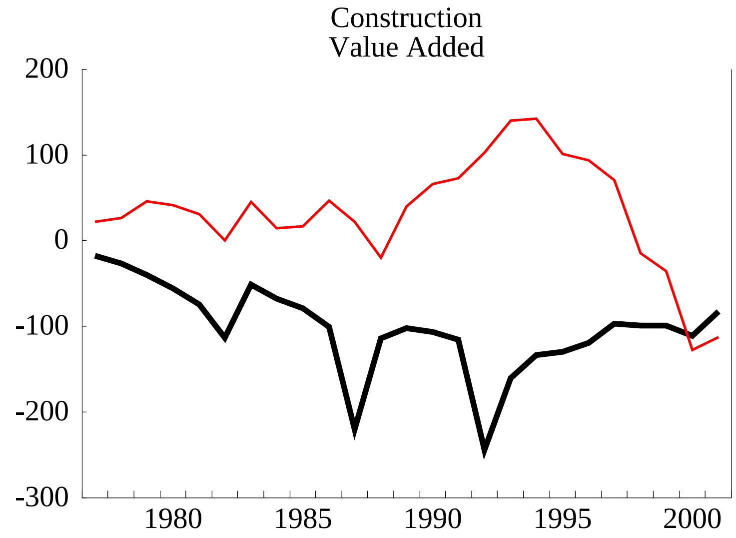
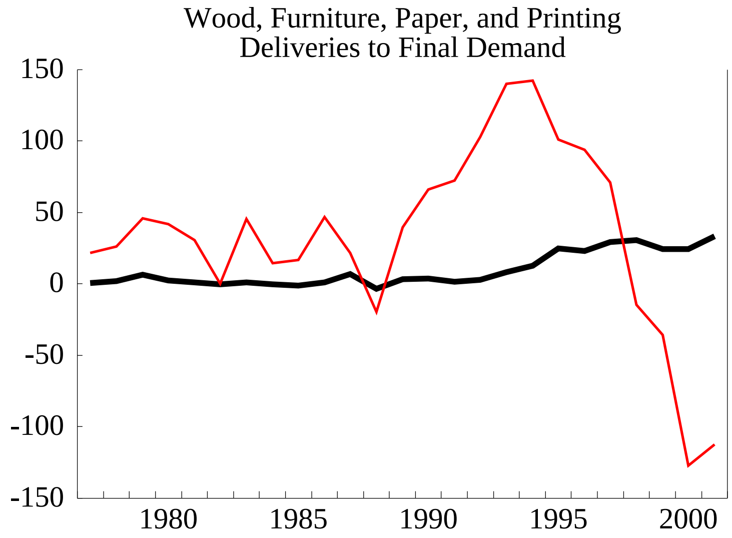


Chart 4 (continued)
Statistical Discrepancy by Industry
(Billions of dollars)

Industry Discrepancy in Black (Thick)   Total Discrepancy in Red (Thin)







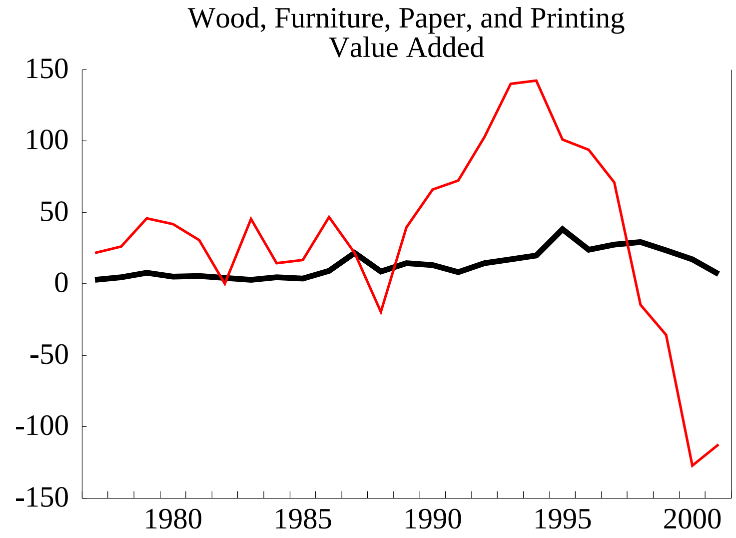


Chart 4 (continued)
Statistical Discrepancy by Industry
(Billions of dollars)
Industry Discrepancy in Black (Thick) Total Discrepancy in Red (Thin)

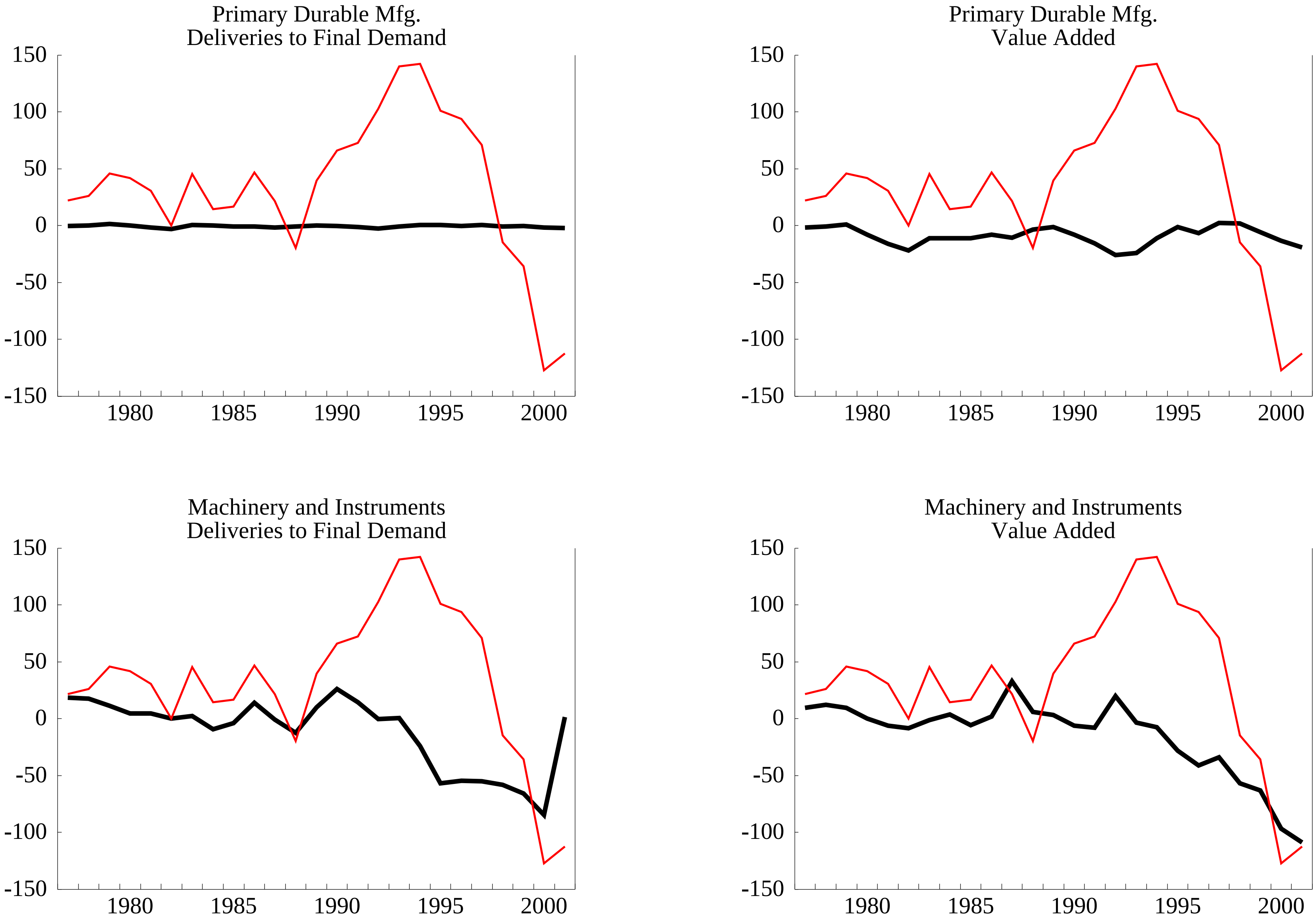
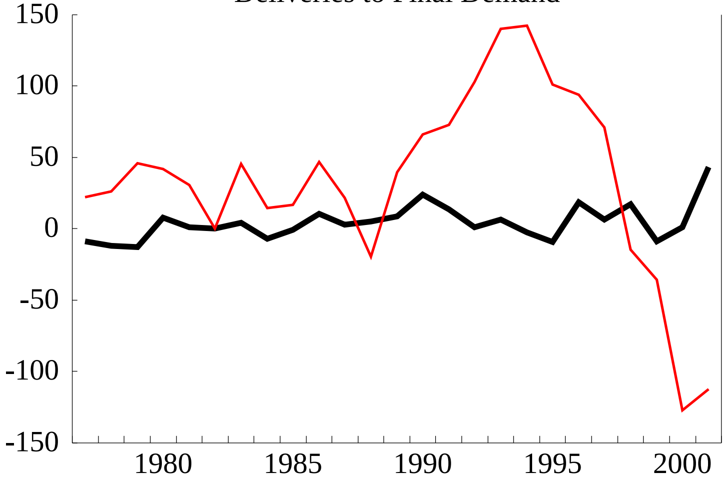


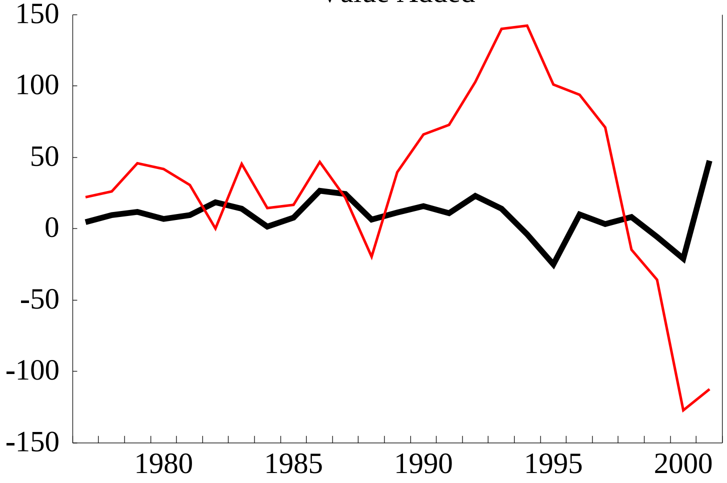
Chart 4 (continued)
Statistical Discrepancy by Industry
(Billions of dollars)

Industry Discrepancy in Black (Thick) Total Discrepancy in Red (Thin)

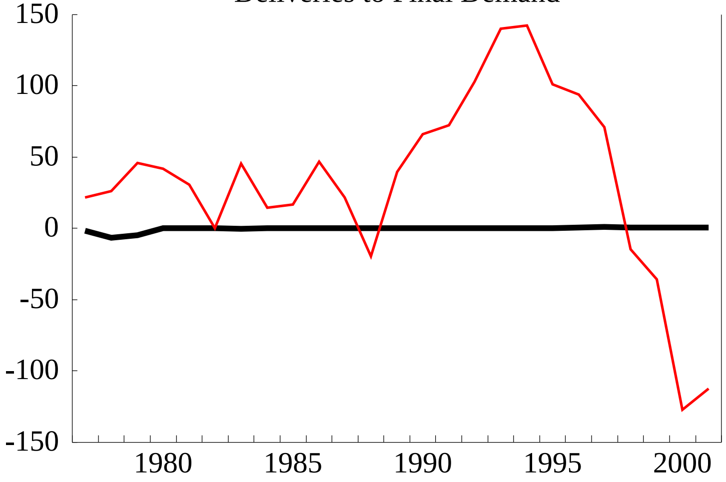
Transportation Equipment
Deliveries to Final Demand



Transportation Equipment
Value Added



Food and Tobacco
Deliveries to Final Demand



Food and Tobacco
Value Added

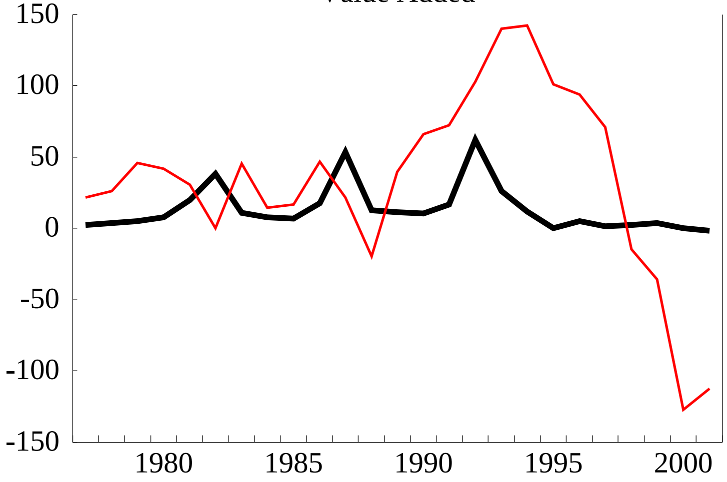
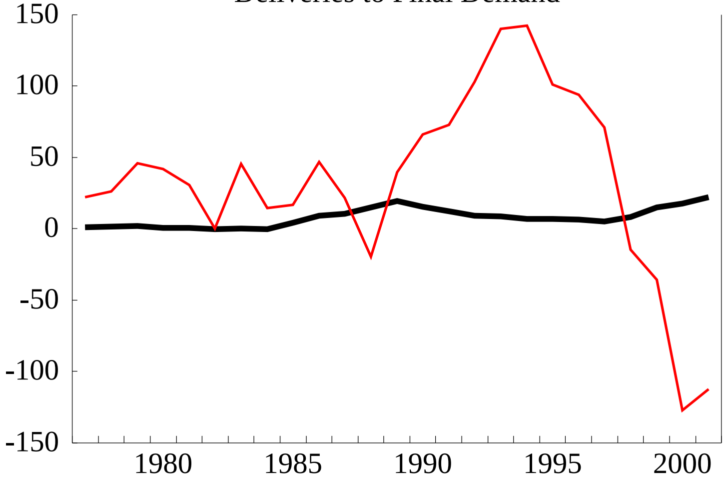


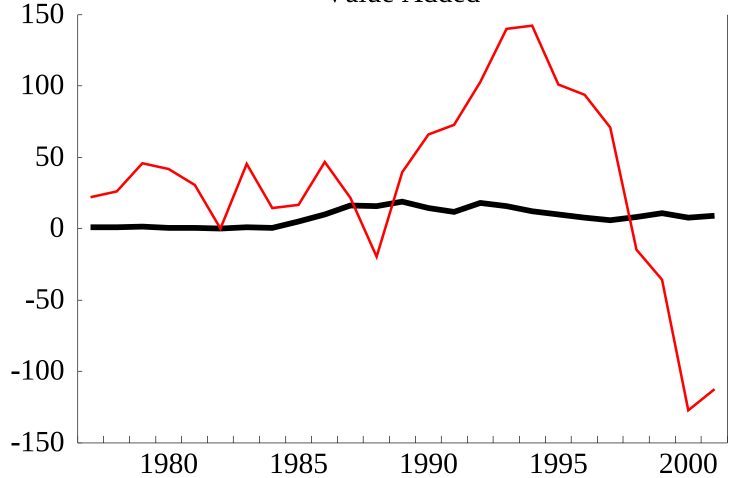
Chart 4 (continued)
Statistical Discrepancy by Industry
(Billions of dollars)

Industry Discrepancy in Black (Thick) Total Discrepancy in Red (Thin)

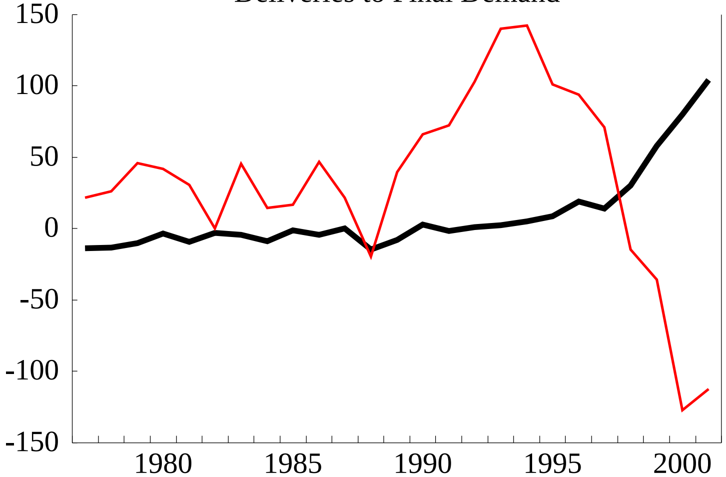
Textiles, Apparel, and Leather
Deliveries to Final Demand



Textiles, Apparel, and Leather
Value Added



Chemicals, Refining, and Rubber & Plastics
Deliveries to Final Demand



Chemicals, Refining, and Rubber & Plastics
Value Added

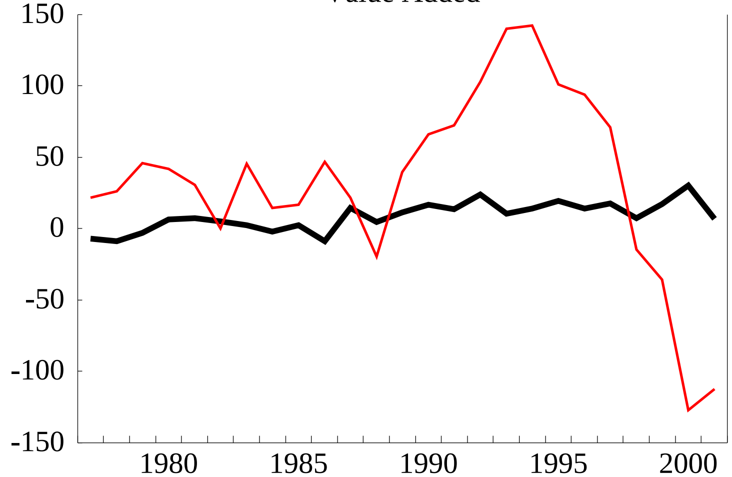
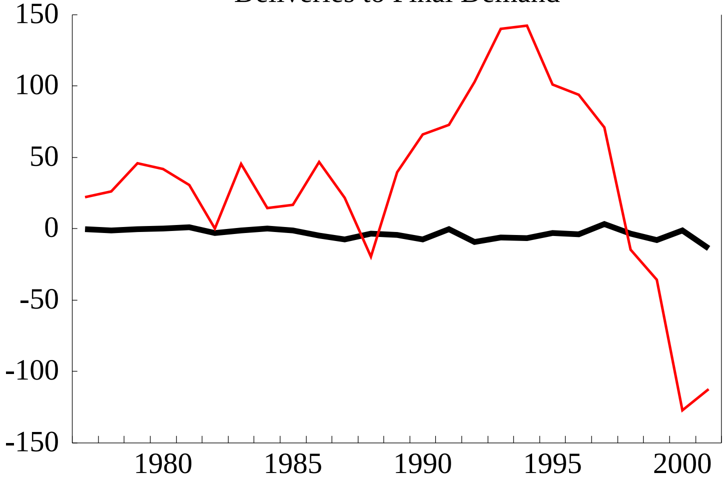
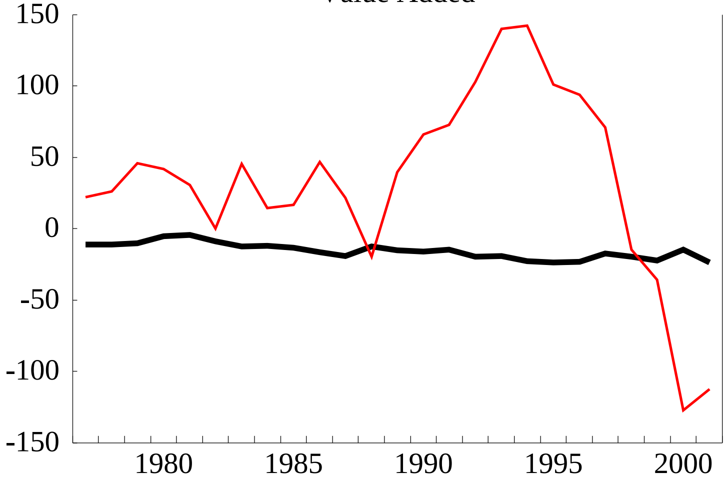


Chart 4 (continued)
Statistical Discrepancy by Industry
(Billions of dollars)
Industry Discrepancy in Black (Thick)   Total Discrepancy in Red (Thin)

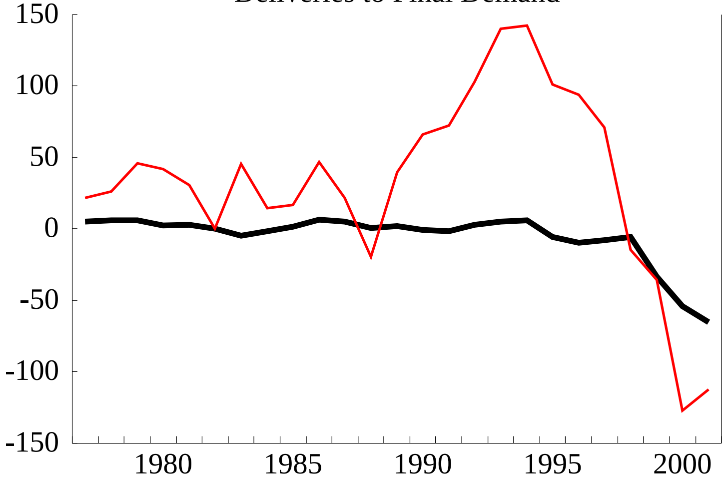
Transportation
Deliveries to Final Demand



Transportation
Value Added



Communications
Deliveries to Final Demand



Communications
Value Added

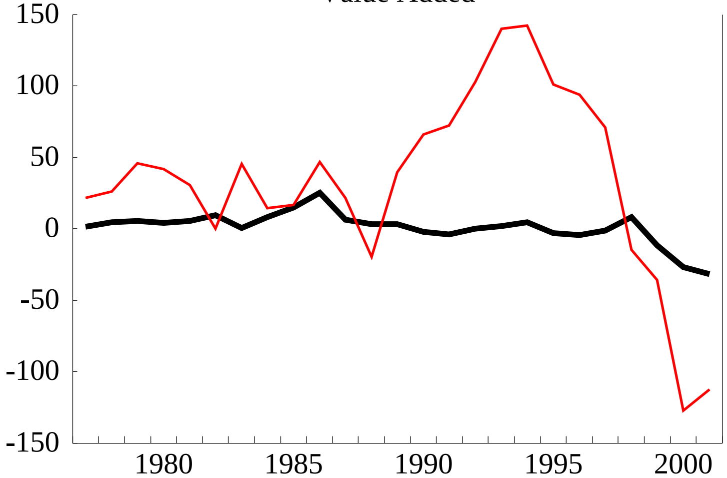


Chart 4 (continued)

Statistical Discrepancy by Industry

(Billions of dollars)

Industry Discrepancy in Black (Thick) Total Discrepancy in Red (Thin)

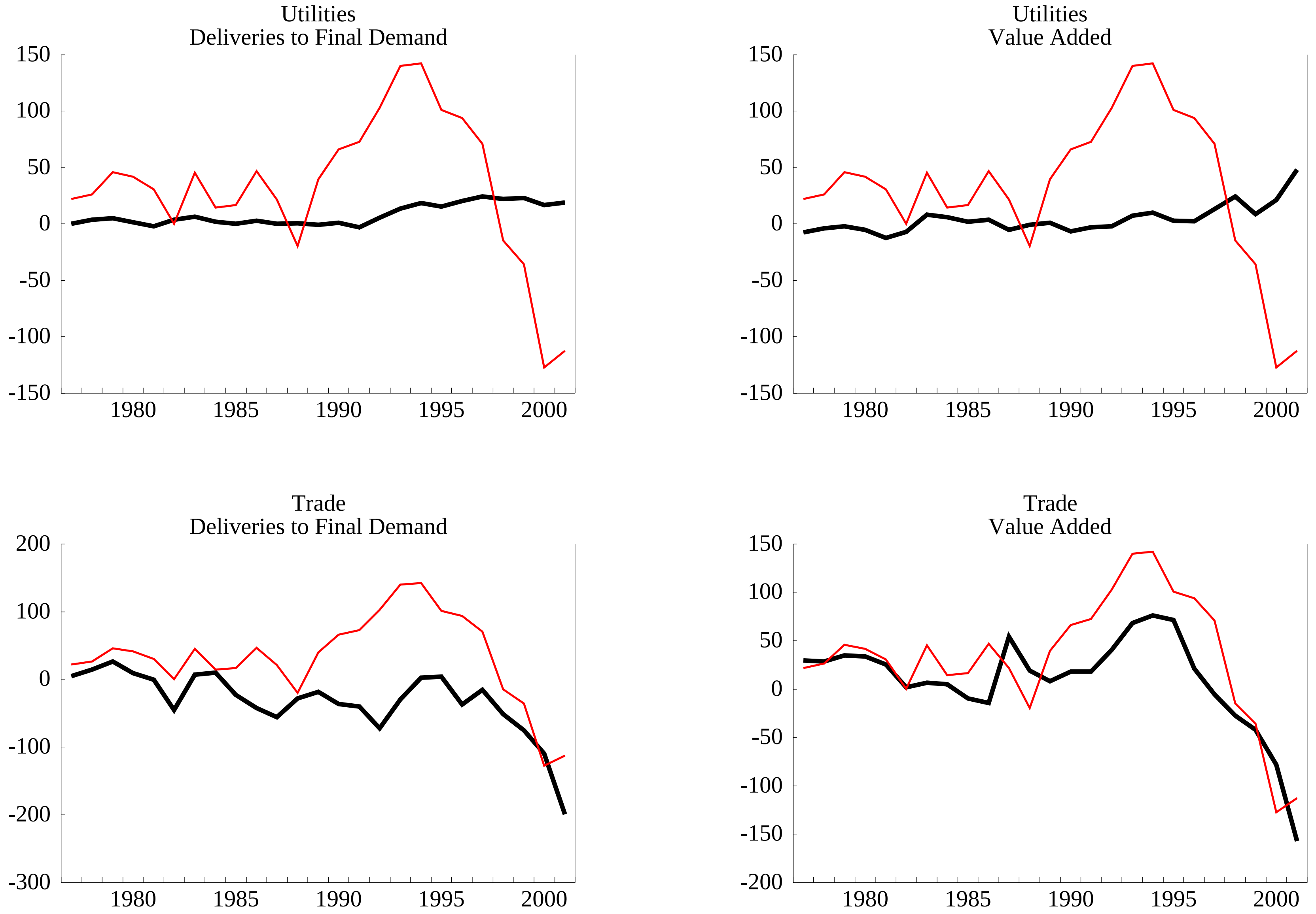
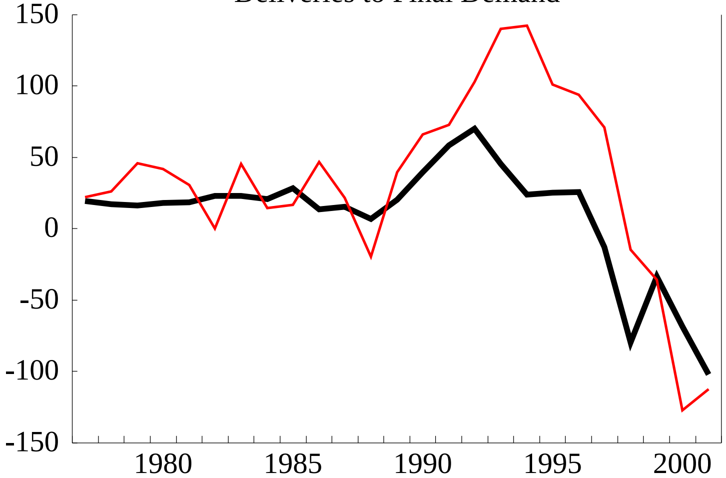


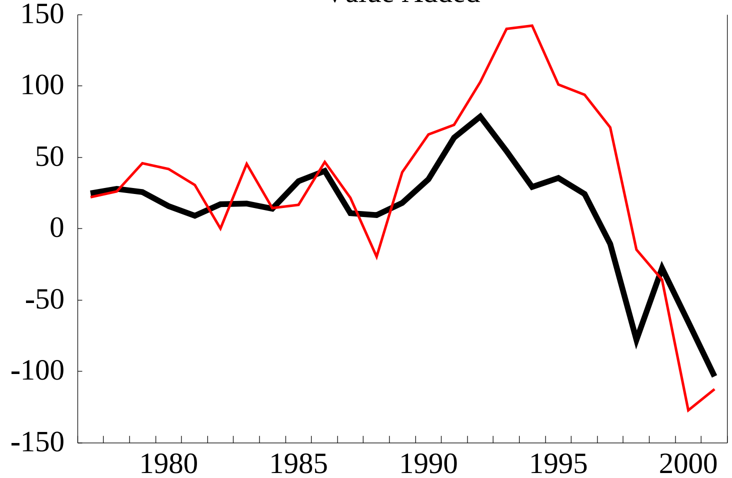
Chart 4 (continued)
Statistical Discrepancy by Industry
(Billions of dollars)

Industry Discrepancy in Black (Thick) Total Discrepancy in Red (Thin)

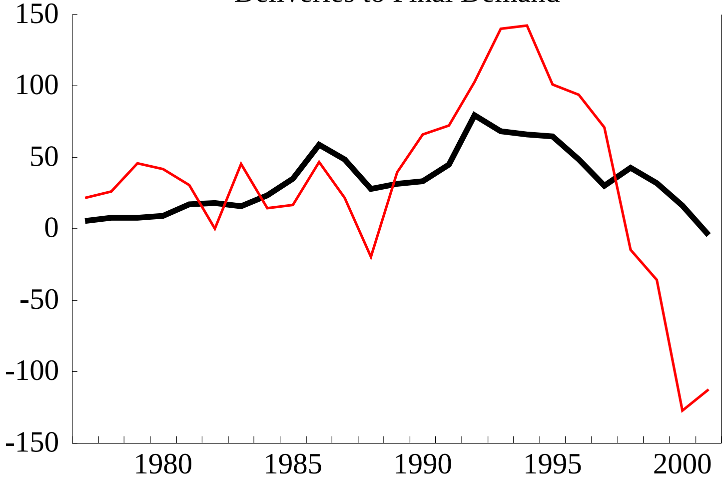
Finance and Insurance
Deliveries to Final Demand



Finance and Insurance
Value Added



Real Estate
Deliveries to Final Demand



Real Estate
Value Added

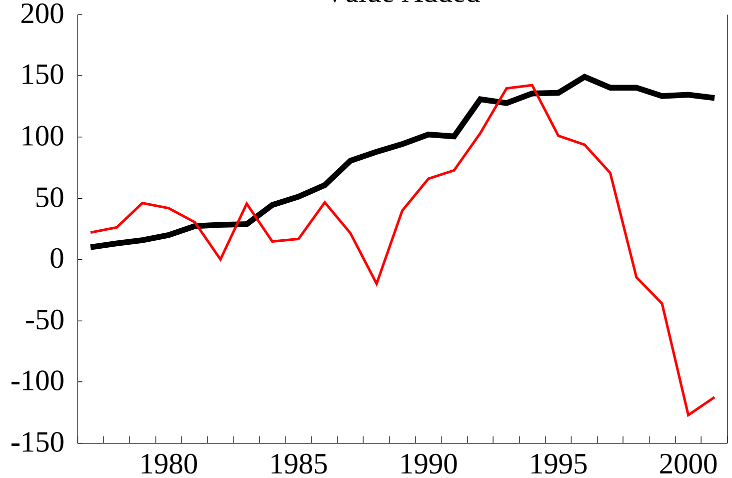


Chart 4 (continued)
Statistical Discrepancy by Industry
(Billions of dollars)
Industry Discrepancy in Black (Thick) Total Discrepancy in Red (Thin)

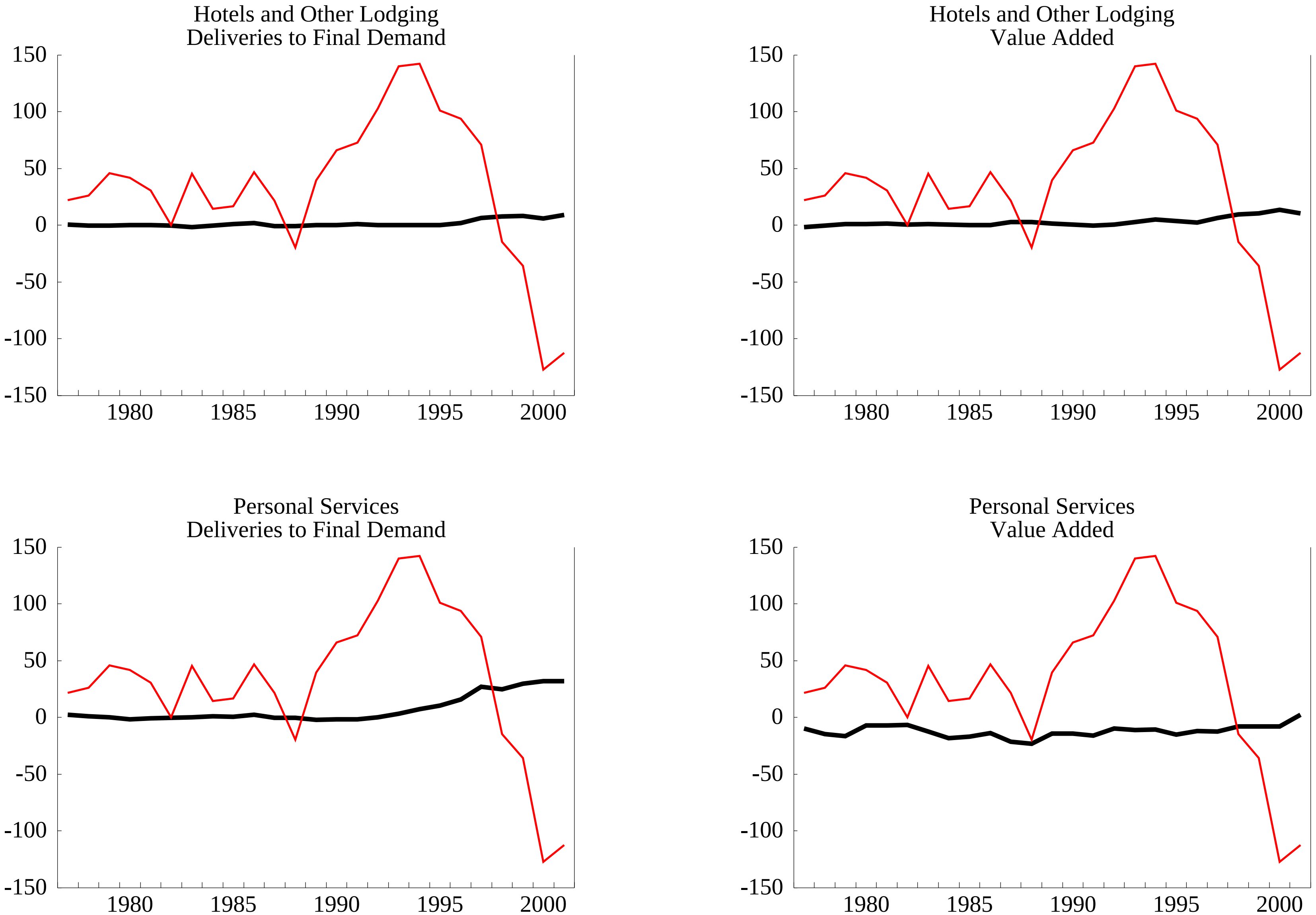
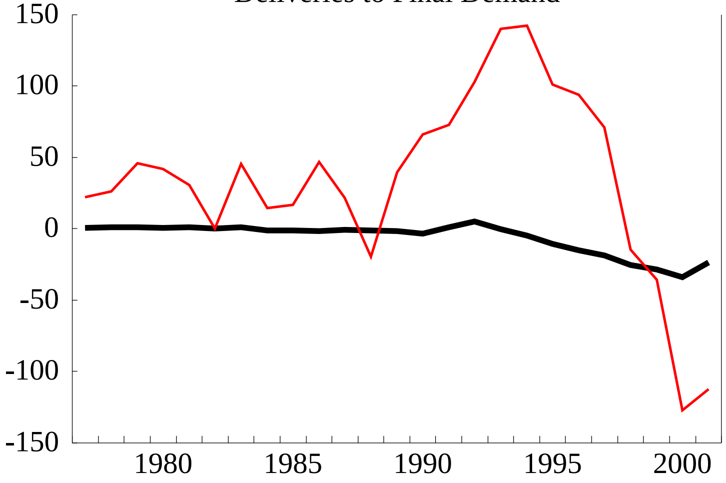
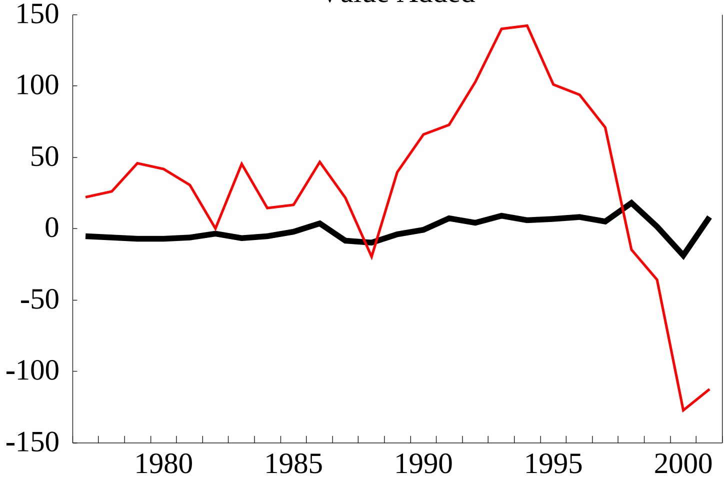


Chart 4 (continued)
Statistical Discrepancy by Industry
(Billions of dollars)
Industry Discrepancy in Black (Thick)   Total Discrepancy in Red (Thin)

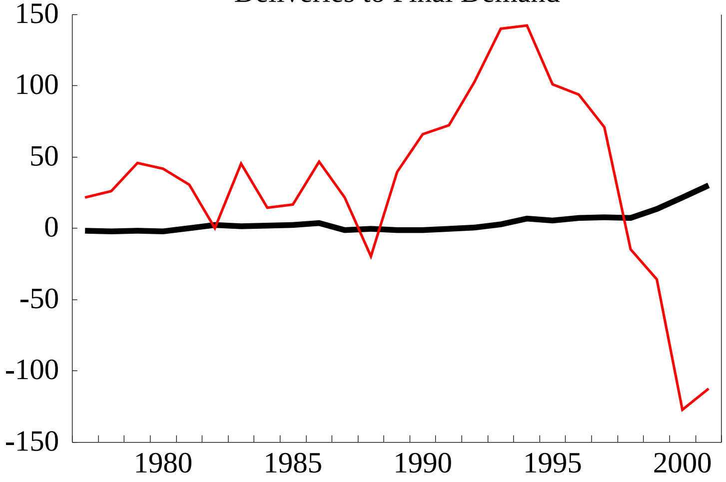
Business Services
Deliveries to Final Demand



Business Services
Value Added



Movies and Recreation Services
Deliveries to Final Demand



Movies and Recreation Services
Value Added

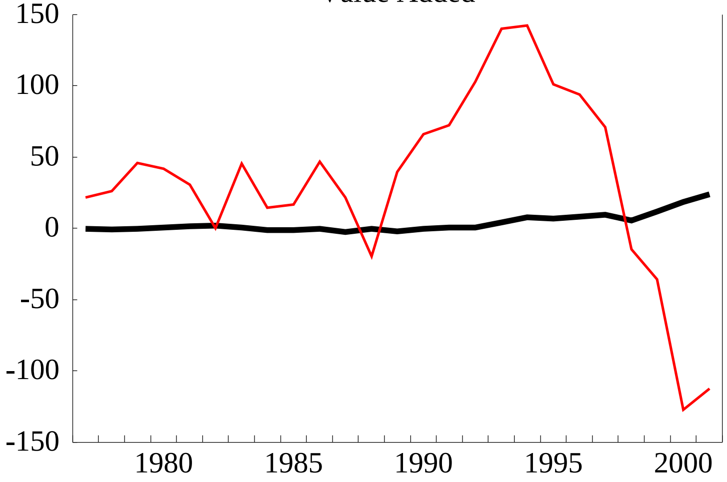


Chart 4 (continued)
Statistical Discrepancy by Industry
(Billions of dollars)
Industry Discrepancy in Black (Thick)   Total Discrepancy in Red (Thin)

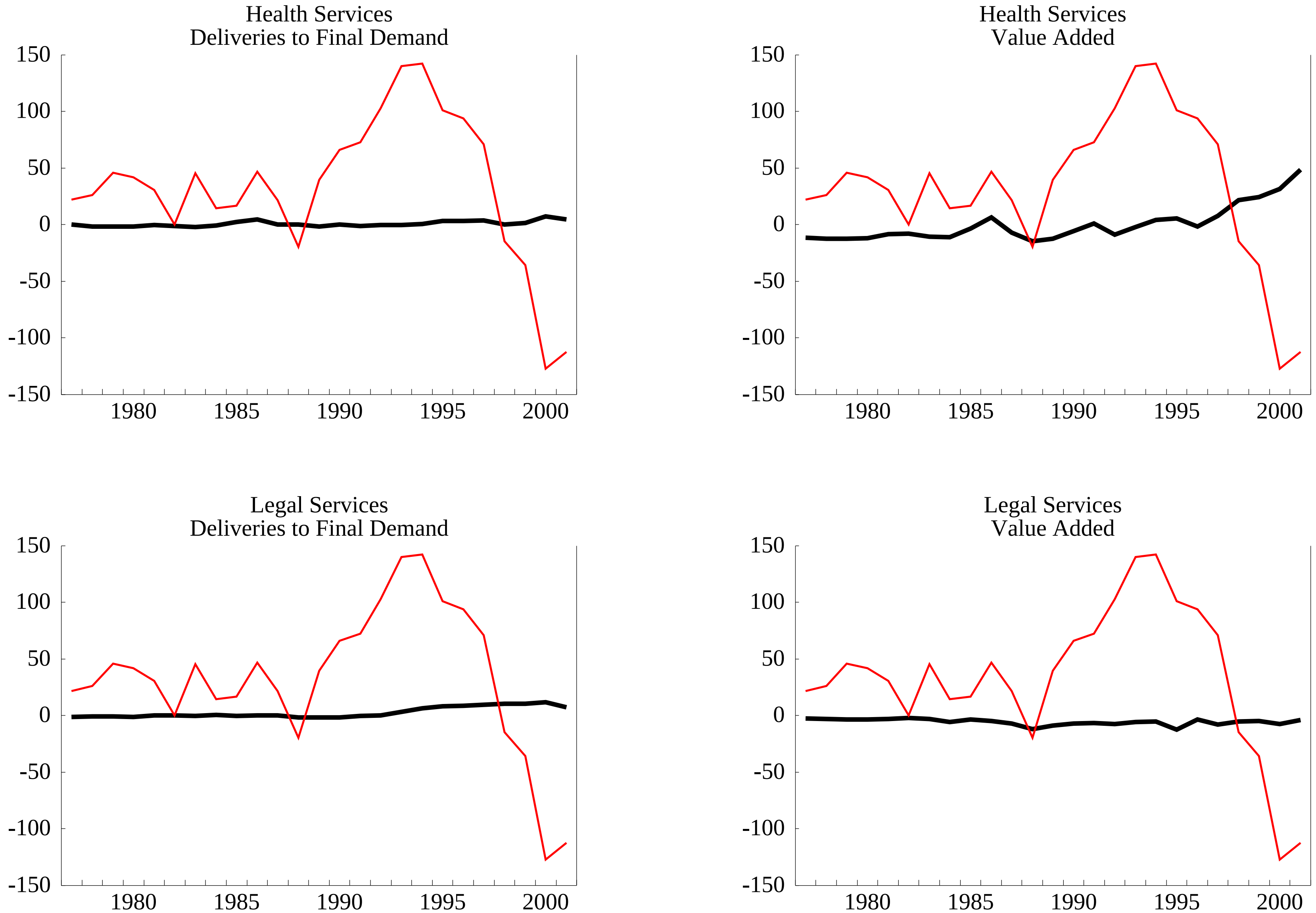


Chart 4 (continued)
Statistical Discrepancy by Industry
(Billions of dollars)
Industry Discrepancy in Black (Thick)   Total Discrepancy in Red (Thin)

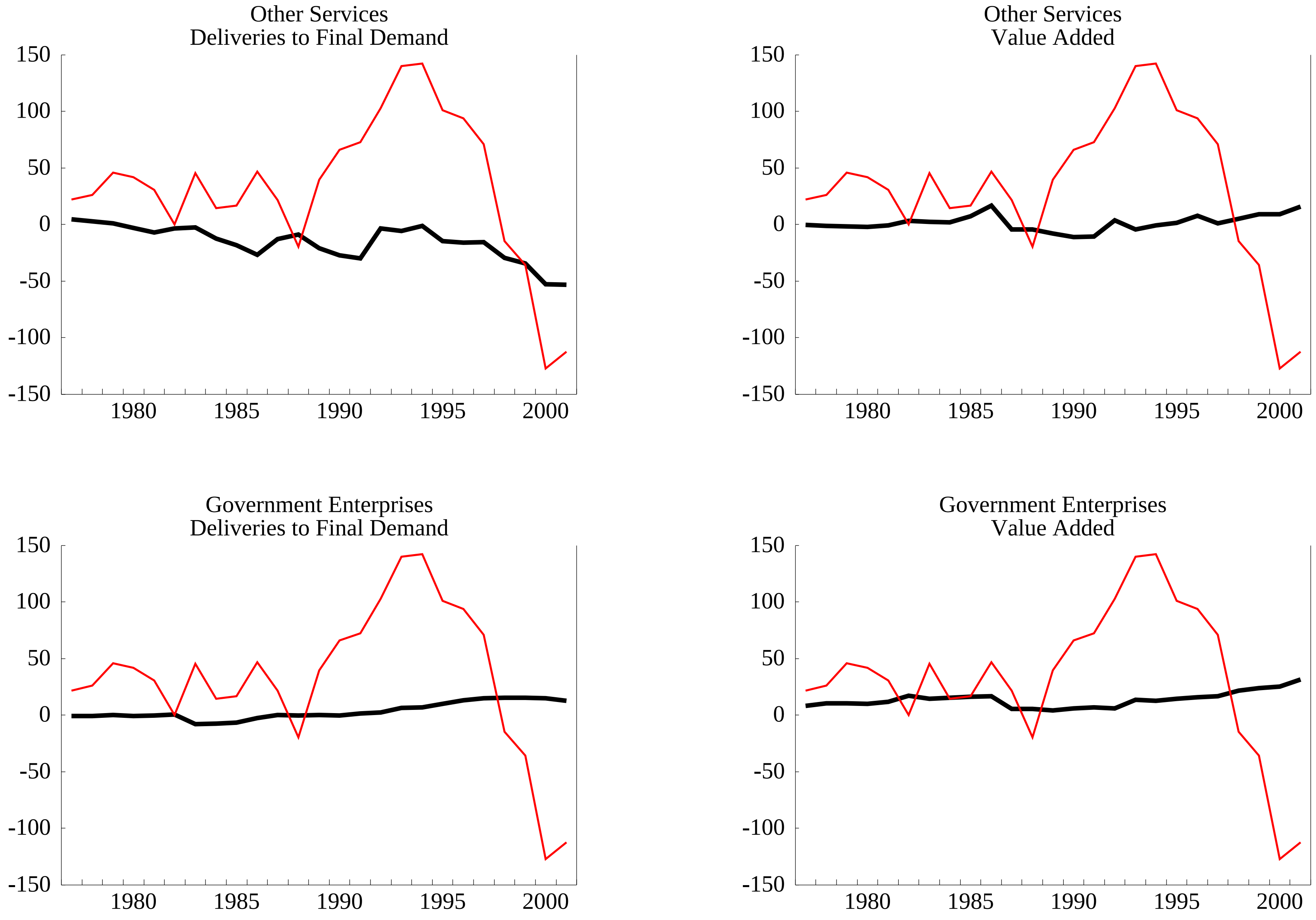
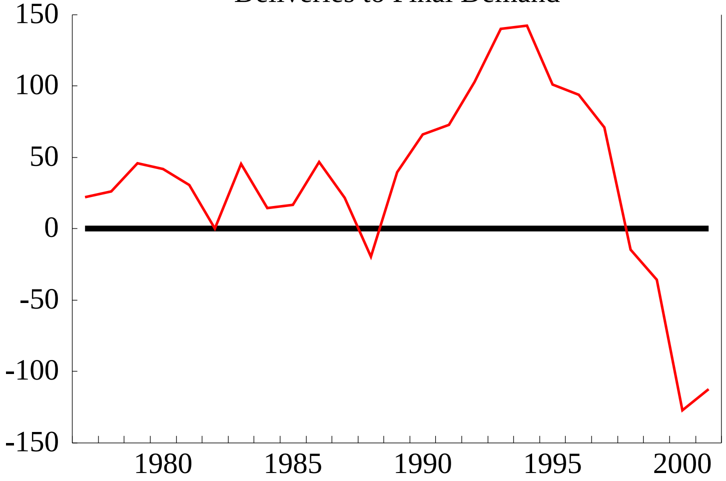


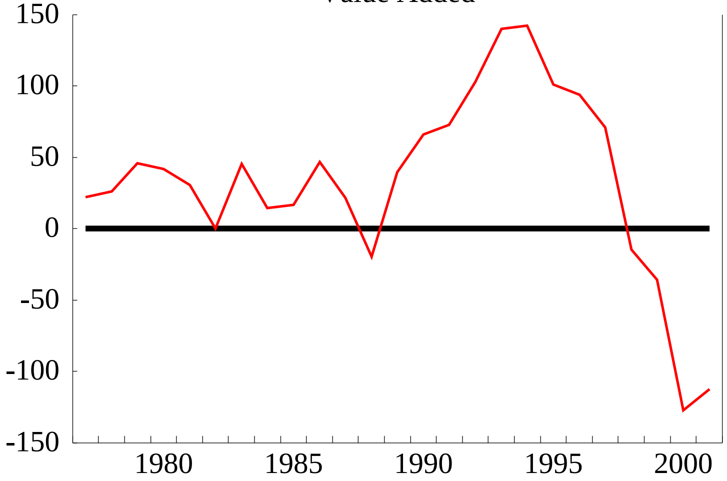
Chart 4 (continued)
Statistical Discrepancy by Industry
(Billions of dollars)

Industry Discrepancy in Black (Thick)   Total Discrepancy in Red (Thin)

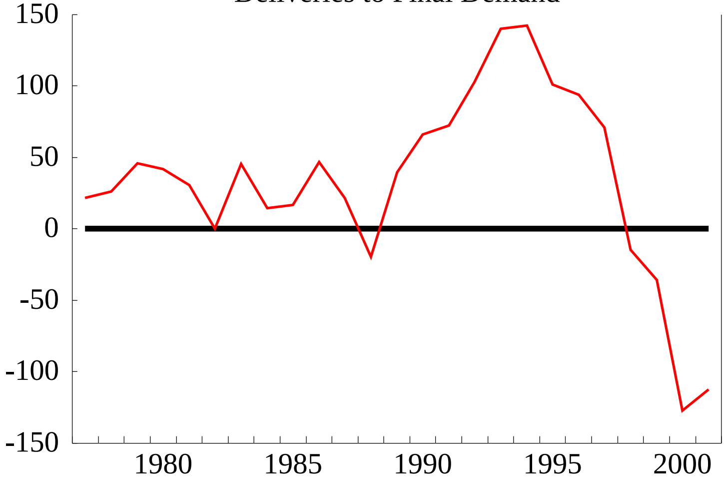
Miscellaneous Industries
Deliveries to Final Demand



Miscellaneous Industries
Value Added



Not Domestic Production
Deliveries to Final Demand



Not Domestic Production
Value Added

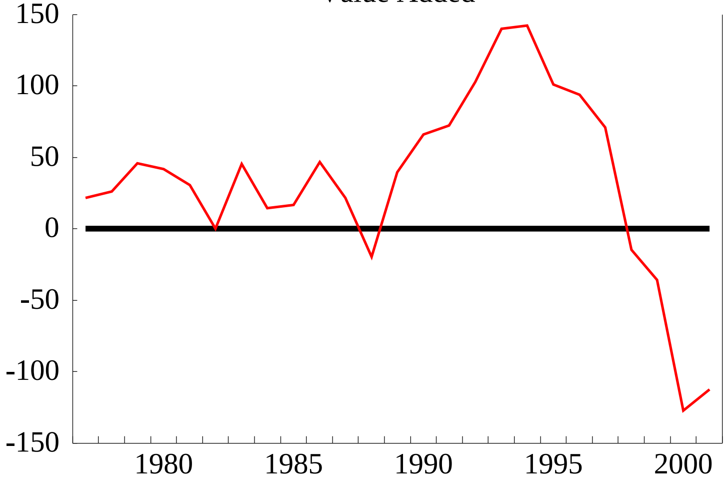
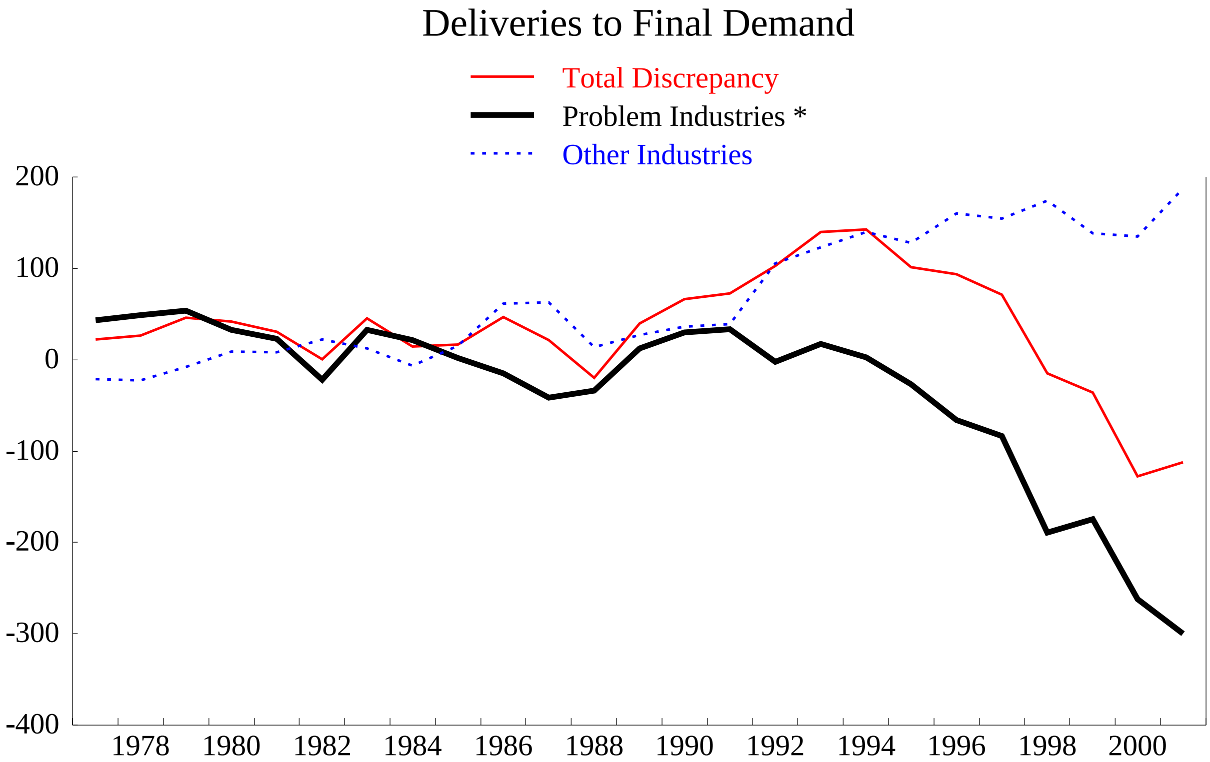
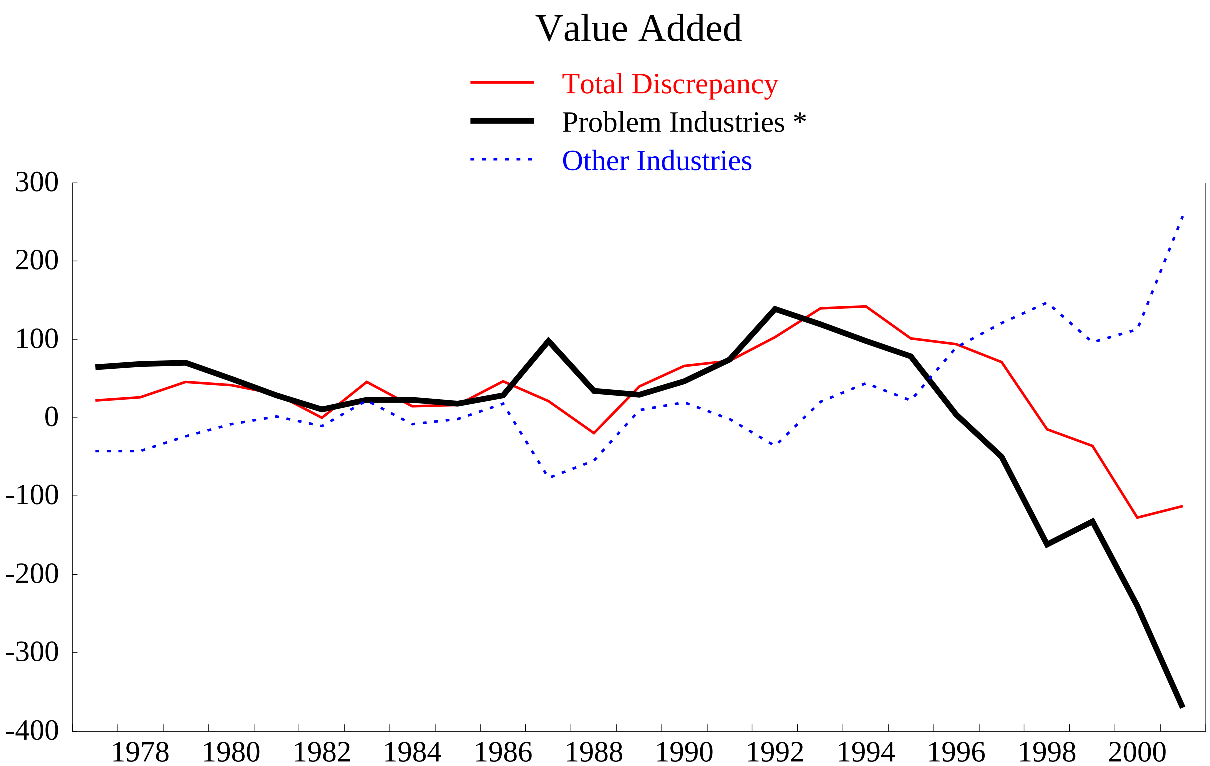


Chart 5
Statistical Discrepancy of Problem Industries
(Billions of dollars)





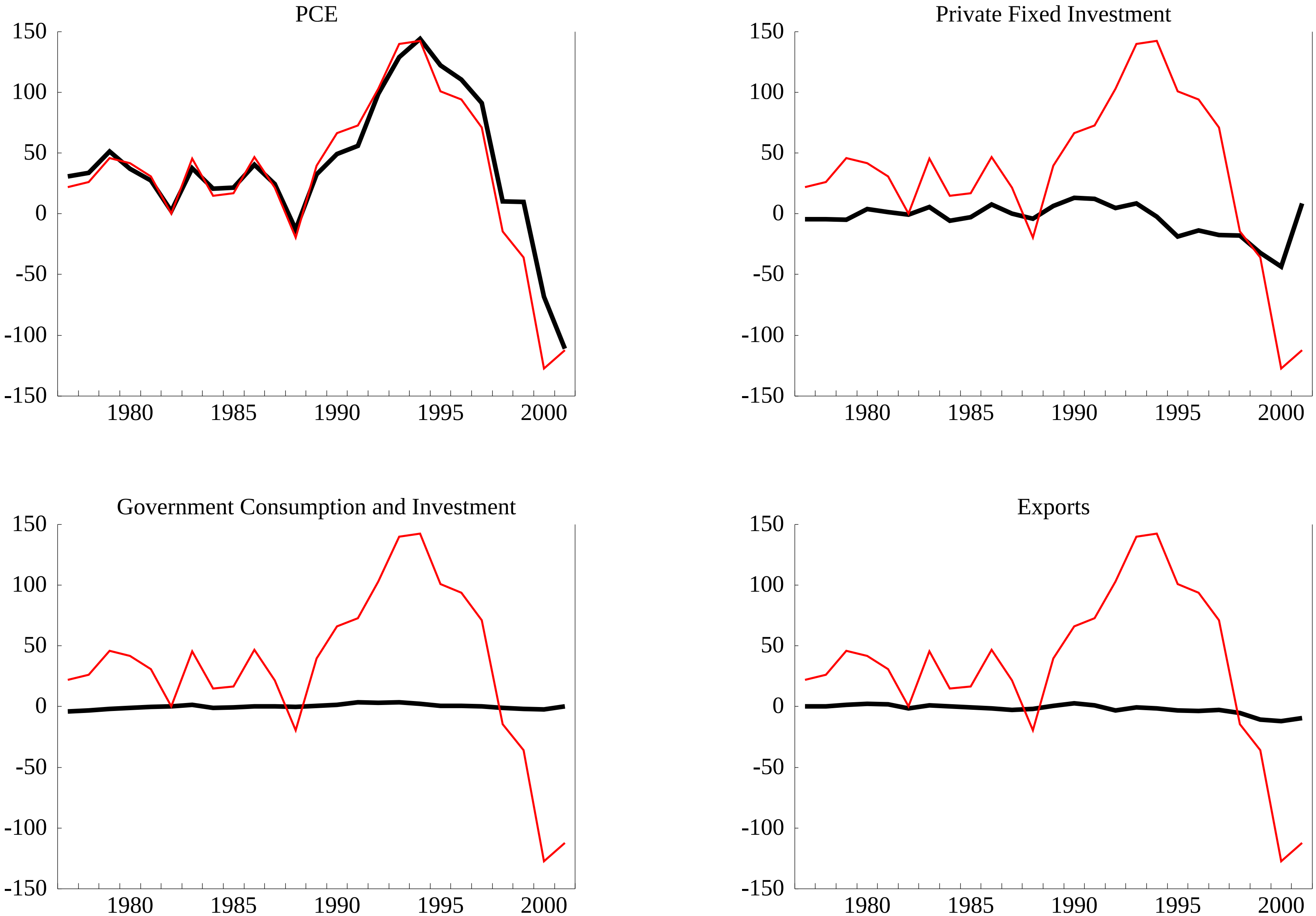
* Problem industries are: Machinery and Instruments, Trade, and Finance and Insurance.

Chart 6

Statistical Discrepancy by Expenditure Category*

(Billions of dollars)

Industry Discrepancy in Black (Thick)   Total Discrepancy in Red (Thin)



*Differences in Inventories are less than $1 billion in any year; differences in imports are restricted to be zero.

Chart 7
Comparison of I-O Systems to Economic Theory
(Bottom axis: Tuning Parameter for Final Demand / Tuning Parameter for Value Added)