SUPPLEMENTAL INFORMATION

A CRITIQUE OF SAUNDERS’ ‘HISTORICAL EVIDENCE FOR ENERGY EFFICIENCY REBOUND IN 30 US SECTORS’

- Danny Cullenward (dcullenward@berkeley.edu)
- Jonathan G. Koomey (jgkoomey@stanford.edu)

Tracing the Source of the Jorgenson Data

This appendix reviews the ultimate sources of the data presented in Jorgenson (2007),¹ which are provided for 35 sectors of the U.S. economy over the period 1960-2005 and upon which the analysis in Saunders (2013)² rests.

As discussed in the main text, the instruction sheet accompanying the Jorgenson dataset specifies three references: Jorgenson and Stiroh (2000),³ Jorgenson (1990),⁴ and Jorgenson, Gollop, and Fraumeni (1987).⁵ After reviewing the instruction sheet, we begin with the most recent of these references and proceed in reverse chronological order, supplementing the available information with additional cross-referenced papers as appropriate. In particular, we resolve ambiguities remaining after reviewing these references by relying on a recent book published after Saunders conducted his study, to which Profes-

---


sor Jorgenson referred us for the most complete description of his methodology (Jorgenson et al., 2013). For the most concise overview of Jorgenson’s methodologies, see Appendix B of this book, which we review in Section 6 here. Although reference to this material is the quickest way to understand the conceptual and empirical issues we raise, our focus in this supplemental information document is on tracing the information available to Saunders, based on the documentation provided in the Jorgenson data at the time of Saunders’ study. This analysis indicates that the major conceptual errors identified in the main text of our paper were (or should have been) ascertainable to Saunders at the time he conducted his analysis.

1. Data Documentation

The instruction sheet accompanying the Jorgenson data provides only a brief description of the underlying data sources, highlighting the data’s mixed origins and inconsistencies with previously reported methodologies:

“These numbers are based on a combination of industry data from the BEA and BLS, and therefore value-added numbers will not match the official NIPA value-added numbers by sector. In addition, the methodology used to estimate real capital and labor input series was changed to better match the BEA data, so that these series do not precisely match those originally reported in Jorgenson and Stiroh (2000).”

There is no additional information in this document to explain the differences with respect to the data and methodologies reported in Jorgenson and Stiroh (2000). Although the documentation provides little other guidance about the source of the data, it clarifies that the production function used to generate the data defines energy commodities as arising from only five industries: Coal Mining, Oil & Gas Mining, Petroleum and Coal Products, Electric Utilities, and Gas Utilities. Thus, the monetary value of commodity flows recorded in the input-output (“I-O”) tables from these sectors will be counted as interme-

---

7 Id. at 437-532; in particular, see id. at 451-58 (describing the primary sources and basic construction of the 1960-2005 data series; see also id. at 461-66 (describing the origins of the data for the five energy-producing sectors in the Jorgenson data).
diately energy inputs in the production function of all sectors.


The most recent reference cited in the data documentation provides more guidance, but falls short of specifying the exact nature of the original data sources:

“Our primary data include a set of interindustry transactions accounts developed by the Employment Projections Office at the BLS. These data cover a relatively short time period from 1977 to 1995. We linked the BLS estimates to industry-level estimates going back to 1958, described by Stiroh, and extrapolated to 1996 using current BLS and BEA industry data. This generated a time series for the period 1958-96 for thirty-seven industries, at roughly the two-digit Standard Industrial Classification (SIC) level.”

The authors continue:

“Our primary data are annual time series of interindustry transactions in current and constant prices, including final demands by commodity, investment and labor inputs by industry, and output by industry …. These data report intermediate inputs and total value added (the sum of capital and labor inputs and taxes) for 185 industries from 1977 to 1995. A major advantage of these BLS interindustry data is that they provide the necessary interpolations between benchmark years.”

Thus, although the exact file or publication used by Jorgenson and Stiroh is not specified, the authors indicate that its format and content is that of a typical I-O matrix. Furthermore, they acknowledge that the BLS interpolated data for the years between benchmark surveys. Turning to methodology, the authors continue:

“We aggregate the data from the BLS’s ‘Make’ and ‘Use’ tables to generate interindustry transactions for thirty-five private business industries at approximately the two-digit SIC level. These tables enable us to generate growth rates of industry outputs, growth rates of intermediate inputs, and shares of intermediate inputs … They also provide control totals for value added in each industry, as the sum of the values of capital and labor services and taxes.”

As this passage shows, although Jorgenson and Stiroh (2000) identify a primary data source—a BLS series—they do not specify the particular reference. In addition, the au-

---


11 Jorgenson and Stiroh (2000), supra note 3 at 207.

12 Id.
thors indicate that they link, extend, and possibly modify these data with other BLS and BEA sources. Finally, this reference does not fully explain the origin of energy price and consumption data. Those methods are described elsewhere.

2.1 Stiroh (1998)

Although Jorgenson and Stiroh (2000) cite this paper to describe how they extended their core BLS data spanning 1977-1995 back to 1960, Stiroh (1998) does not actually explain how this happened. Stiroh (1998) merely references Jorgenson et al. (1987) and says that the data he used were updated through 1991.13


Elsewhere, Jorgenson (1990) clarifies the exact source of one of the BLS data series. Unfortunately, the original source cited in 1990 is listed as a “computer tape,”14 suggesting that the original data file is not readily available. When we contacted BLS representatives, they were unable to identify this file without undertaking significant archival research. It may be in the Bureau’s archives, but even if so, the data series has been discontinued. Thus, tracing the exact origin of the data requires additional assumptions.

Currently, the BLS website for inter-industry data clarifies that its annual I-O data are derived from BEA’s Industry Economic Accounts.15 However, BLS provides data only for the period 1993–2012,16 which does not extend as far back as 1977, the beginning

13 Stiroh (1998), supra note 10 at 117.  
16 Id. (offering I-O data for the period 1993-2012) (as of July 23, 2014).
date of the BLS data referenced in Jorgenson and Stiroh (2000). This suggests that the currently available inter-industry data are distinct from the original Jorgenson source. Although BLS no longer publishes I-O data for years prior to 1993, as it discontinued its annual I-O SIC data series, BEA’s historical I-O SIC data extend back to 1977. Because the computer tape referenced by Jorgenson (1990) was not readily available, BLS representatives could not confirm whether or not it was based on historical BEA data. Given that BLS has traditionally provided its own interpolations of annual I-O data between the BEA benchmark I-O survey years, however, it seems very likely that the computer tape upon which Jorgenson (1990) did as well.


This book provides the earliest description of Jorgenson’s methodology and data sources. It is the foundational work upon which subsequent papers and books are based. Although its period of study (1948-1979) is different than all of the other references examined here, the core methodology appears consistent with that of subsequent studies. In turn, Jorgenson et al. (1987) provide a fairly complete explanation of the underlying data used in that study. Additionally, the book offers a road map to the issues associated with linking the various data sources, as well as some of the activity redefinitions between different SIC classification vintages. While these methodological descriptions are helpful, they do not apply directly to the published 2007 Jorgenson data, which use different definitions for industrial categories, as well as significantly different primary data sources.


18 See the discussion in Section 6, infra.

19 Astute readers may wonder why Jorgenson relied on BLS data. If BLS was merely offering its own interpolation of the years between BEA benchmark surveys, why not use BEA’s own annual I-O data series? Here we can only speculate; in informal private conversations, government economists have suggested that prior to the 2009 comprehensive revisions, BEA’s annual I-O tables could not be reconciled perfectly with the BEA benchmark I-O tables. If true, this would provide reason for Jorgenson to prefer BLS over BEA I-O data.

that are needed to extend the published data to the year 2005.\textsuperscript{20}

Briefly, Jorgenson et al. (1987) rely on three data sources: (1) inter-industry data from the BLS covering 1959-1979,\textsuperscript{21} adjusted for changes in national income accounting conventions; (2) gross output and inter-industry data from the BEA covering 1948-1958;\textsuperscript{22} and (3) a series of consulting reports from Jack Faucett Associates covering nonmanufacturing output for industries for the period 1948-1958,\textsuperscript{23} supplemented by unpublished BEA data for six missing industries, with extrapolation for missing years.\textsuperscript{24} The extensive modifications required to integrate and harmonize these data sources are summarized over several pages.\textsuperscript{25}

Note that Jorgenson and Stiroh (2000) cites Stiroh (1998) to explain the source of the data going back to 1960. Stiroh (1998) in turn cites Jorgenson et al. (1987). This strongly suggests that the BLS data referenced in the paragraph above are used to extend whatever other BLS data exist back to 1960. As discussed in Section 3 above, however, BLS no longer publishes these older data series, so it was not possible to review the primary

\textsuperscript{20} The correspondence between the 45 sectors studied in this book and the 1972 SIC classifications is found in Table 5.1 in Jorgenson et al. (1987), supra note 5 at 153-55. Note that the 1972 SIC classification is different from the 1987 SIC classification, which was the last re-definition of SIC prior to the introduction of NAICS. Note also that the Jorgenson data released in 2007 include 35 sectors, so the correspondence in the older study is merely indicative of a general problem that remains in the later Jorgenson data.

\textsuperscript{21} The reference in Jorgenson et al. (1987) is to BLS (1979), but the book’s reference list is unclear as to which citation was intended. The two options are (1) BLS (1979a), Capital stock estimates for input-output industries: Methods and data, Bulletin 2034; and (2) BLS (1979b), Time series data for input-output industries, Bulletin 2018. Neither source is currently available from BLS, although it is possible that the BLS archives contain them.


\textsuperscript{23} Jack Faucett Associates (1975), Output and employment for input-output sectors. Report for the Bureau of Labor Statistics. Note that this report covers data only for years outside of the range presented in the final 2007 Jorgenson data, so it presumably is not relevant for our purposes.

\textsuperscript{24} Jorgenson et al. (1987), supra 5 at 156-57.

\textsuperscript{25} Id. at 149-51.
In addition to providing more guidance about the underlying data in this earlier study, Jorgenson et al. (1987) describe their basic methodological approach, which is continually referenced by subsequent studies, including the documentation accompanying the Jorgenson data. The Jorgenson et al. (1987) method for addressing intermediate inputs takes estimates of gross output and value added by industry as a starting point. For each industry, total intermediate input is defined as gross output minus value added, with all terms expressed in current prices. Next, Jorgenson et al. transform the inter-industry I-O tables as follows:

“Our first step … is to develop an interindustry matrix in current prices for each year, 1948-1979, that conform to national income accounting definitions. The eleven 80-order [BEA] interindustry benchmark matrices are the basis for determining the share of intermediate input used by each of the forty-five industries for which we measure intermediate input. Household and government industries are excluded. The 1963 367-order interindustry matrix was used to provide additional detail … We interpolated the shares between benchmark years. The 1972 shares were used for all subsequent years. At this point we have a complete set of interindustry matrices, 1948-1979, with information on current-price intermediate inputs, by sector of origin and destination, on a basis conforming to national income accounting conventions.”

The authors then describe how they generate equivalent I-O matrices in constant prices:

“We employ output deflators in purchasers’ prices for each industry … It is important to note that the output deflator used in evaluating the value of a sector’s output is not the same as the deflator used in evaluating that sector’s deliveries to intermediate demand. The former is measured in producers’ prices; the latter is measured in purchasers’ prices.”

These price indices are taken from the BEA and Jack Faucett Associates data on gross

---

26 These older data series appear to be individual computer files, rather than published reports that would be available in research libraries or in general federal government archives.

27 Note that this particular reference groups all material, energy, and service inputs into one aggregate intermediate input category—unlike the 2007 Jorgenson data, in which energy inputs (“E”) are treated separately from other intermediate inputs (“M”).

28 Jorgenson et al. (1987), supra note 5 at 159.

29 Id. at 159-60.

30 Id. at 160.
industry output, implying that price indices are not sector-specific—i.e., that there is one price per commodity, a price that all purchasers face.

Although this key assumption is never discussed explicitly in Jorgenson et al. (1987), it is strongly suggested by the origin of the data for the price indices. Unless the BEA and Jack Faucett Associates data series actually measure the price paid for every inter-industry transaction category, some sort of assumption would be necessary to assess the value and flow of all inter-industry transactions. Because the level of data required to measure the prices of all inter-industry transactions would be equal to that required to measure aggregate inter-industry expenditures, the fact that there is no mention of the data sourcing or integration challenges supports the conclusion that the BEA and Jack Faucett Associates data provide economy-wide price indices for each industry’s output. Therefore, this interpretation is consistent with the methodology described in Jorgenson et al. (1987). It is also confirmed explicitly in Jorgenson et al. (2005) and Jorgenson et al. (2013), as discussed below. This key assumption can also be confirmed by direct examination of the Jorgenson data. Moreover, Professor Jorgenson kindly confirmed that all sectors face the same economy-wide average price for any given sector’s representative commodity in a June 2014 email exchange.

5. Jorgenson, Ho, and Stiroh (2005):

Despite not being cited in the documentation accompanying the Jorgenson data, Jorgenson et al. (2005) contains significant additional information about the data and methods common to Jorgenson’s basic approach. It was also cited by Saunders (2013). Jorgenson et al. (2005) assess the effects of intermediate inputs on economic growth over

\[ \text{id. at 158.} \]

\[ \text{It is quite likely that the mathematical descriptions of Jorgenson’s model make this assumption explicitly, but we did not review all equations in the book to locate the necessary information. Subsequent references make clear this is a feature of Jorgenson’s basic modeling approach, as discussed in Sections 5 and 6, below.} \]

\[ \text{Email correspondence between Professor Dale Jorgenson, Danny Cullenward, and Jonathan G. Koomey (June 10-17, 2014) (on file with authors).} \]


\[ \text{Saunders (2013), supra note 2 at 1318.} \]
the period 1977-2000, explicitly citing the methods developed in Jorgenson et al. (1987) and Jorgenson (1990).

Most importantly for our purposes, Jorgenson and colleagues present a clear explanation of the origin of the intermediate input price data they produce:

“A key assumption is that the price of each intermediate input $i$, $P_i^{X}$, is the same for all purchasing sectors. Each of the 41 business industries is an aggregate of many sub-industries classified at a finer level of detail in the input-output accounts. Each input denoted by $X_{ij}$ is therefore a basket of these sub-commodities. Even if the prices of all sub-commodities at the finest level were the same for all buyers, the different shares bought would produce a different $P_i^{X}$ for each purchasing industry. Moreover, prices paid are actually different for buyers in different locations and the official indices are averages over the whole country. Our common input price assumption, therefore, averages out some effects that might be observed in studies of more detailed industries, or in studies using surveys of actual quantities.”

Because Jorgenson et al. (2005) explicitly cite Jorgenson et al. (1987) and Jorgenson (1990) as source of the 2005 book’s methodology, it is quite likely that this assumption is consistent with the approach taken in previous studies. We note also that Saunders (2013) cites Jorgenson et al. (2005), suggesting this point was either understood or could have been understood at the time Saunders completed his analysis. Again, the fact that all sectors face the same price for a given intermediate input can be confirmed directly in the published Jorgenson data.

Expanding prior descriptions of primary data sources, Jorgenson et al. (2005) cite the BEA benchmark I-O tables for 1977, 1982, 1987, and 1992 as the starting point. The authors continue, noting that BLS uses these tables to generate a time series of inter-industry transactions tables for 1983-2000 that cover 192 sectors, although no specific reference is provided. These I-O tables are then combined with estimates of gross output by industry, providing the foundation for the production function Jorgenson and col-

36 Jorgenson et al. (2005), supra note 34 at 96.
37 Id. at 90 (“Our methodology for measuring industry output, intermediate inputs, and value-added was introduced by Jorgenson, Gollop and Fraumeni (1987) and extended by Jorgenson (1990a).”) (citing Jorgenson et al. (1987), supra note 5, and Jorgenson (1990), supra note 4, respectively).
38 Saunders (2013), supra note 2 at 1318.
39 Jorgenson et al. (2005), supra note 34 at 107.
40 Id.
leagues apply. As the authors put it:

“Our primary data source is the gross output data produced by Office of Occupational Statistics and Employment Projections at the Bureau of Labor Statistics, which can be linked with corresponding input-output tables at a relatively narrow classification level. We refer to this data as BLS-EMP data. We do not make use of the gross output data from the BEA GDP by Industry data set (referred to as BEA data) because that data cover a relatively short period beginning in 1987 and include relatively broad industry classifications.”

Finally, Jorgenson et al. (2005) also indicate using an earlier version of BLS data that covers the period 1977-1995. Again, no specific citation is provided, although this time the authors cite to the general BLS website and thank several BLS economists who assisted them in parsing the data. Unfortunately, BLS no longer offers any of these data series on its website.

With respect to the newer BLS data, Jorgenson et al. (2005) report that additional assumptions are required to track Electric and Gas Utilities separately:

“The BLS-EMP industry ‘combined utilities’ is decomposed by allocating a fixed share to Electricity and the remainder to Gas Utilities. The share is taken from the previous version of the BLS-EMP data covering 1977-95, which does not have such an industry.”

This passage indicates that the expenditure flows from two of the five energy-producing sectors are inferred by decomposition from an aggregate category. The documentation does not explain exactly how the shares are assigned to each of the two utility categories, but it does say that the data from which the shares are determined covers the period 1977-1995. That period was just one episode in an era of continuous change for both the commercial and industrial sectors, and is unlikely to provide a representative basis from which the actual shares of electric and gas utility commodities could be accurately projected. Assuming that these two commodities approximate the use of natural gas and retail electricity as fuels, one can compare the share of each fuel in the commercial and industrial energy mix over time and see that each has been changing both before and after the period from which Jorgenson made his extrapolations (Figure SI-1 and Figure SI-2).

---

41 Id. at 89.
42 Id. at 107. The authors cite a BLS website that no longer contains any relevant data.
43 Id. at 109. No additional information is provided as to how these shares are determined.
(In a later book, Jorgenson et al. (2013) clarifies that the combined utility sector was allocated 90% to gas utilities and 10% to electric utilities.\footnote{Jorgenson et al. (2013), supra note 6 at 452.})

In general, Jorgenson et al. (2005) provide more detail than previous studies about how various data sources were linked, as well as the differences between BLS and BEA sources. For example, the authors provide a description of the necessary assumptions and processes to link the two BLS data series covering different periods.\footnote{Jorgenson et al. (2005), supra note 34 at 110.} The authors also compare a variety of alternative data sources to their preferred choices, showing how significant variation is possible depending on the chosen primary data source (see Table SI-1). As these comparisons show, measurement of gross output for energy-producing sectors varies widely, implying that the variation could be even more significant for the I-O tables.

6. \textit{Jorgenson et al. (2013)}:

Jorgenson et al. (2013) confirms that the published Jorgenson data have their origins in the BEA benchmark I-O surveys for 1977, 1982, 1987, and 1992.\footnote{Jorgenson et al. (2013), supra note 6 at 451.} This reference also indicates that surveys for 1997 and 2002 were excluded. The reasoning is as follows. In 1997, the national accounting classification scheme in the United States switched from SIC to NAICS, creating a discontinuity in the national accounts (see Section 7, below). Because the majority of the Jorgenson data period (1960-2005) covers years in which data were collected in the SIC scheme, Jorgenson et al. faced a choice about how to handle the mid-period change in accounting systems. Ultimately, they retained the SIC approach, rather than converting the data to NAICS or bridging the two accounting systems in 1997. Thus, because the BEA benchmark surveys for 1997 and 2002 were published only in NAICS, they were not included in the generation of the Jorgenson data.\footnote{Id. at 451, FN 6.}

Furthermore, Jorgenson et al. (2013) clarifies that the Jorgenson data were created by merging three BLS-EMP I-O data tables—each interpolating the annual data in between
benchmark I-O survey years—for 1960-1985, 1977-1995, and 1983-2000. Although Jorgenson et al. (2013) identify the institutional author (BLS-EMP) and time span of each series, the website they provide no longer contains any historical I-O data series. When one of us (D.C.) reached out to BLS-EMP by phone in August 2013, the BLS economist with whom we spoke was unable to locate any historical I-O data series covering these periods; the oldest BLS I-O data currently published online stretches back only as far as 1993. The BLS economist noted that while BLS-EMP used to create interpolated annual I-O tables, their efforts have been discontinued and are no longer published online. It was possible, the BLS economist told us, that the series were available in the BLS archives, but without more identifying information about the sources in question, the search process would require significant time and expense. We did not pursue the matter any further, having confirmed with BLS-EMP that their previously published annual I-O surveys would have also interpolated the data between benchmark BEA I-O survey years.

Jorgenson et al. (2013) also provides some details on the challenges of merging these older BLS-EMP data series into a consistent final product. For example, the 1983-2000 series covered 192 sectors, each of which had to be mapped to a corresponding sector in the Jorgenson data framework, which covers only 35 more aggregate sectors. The other two BLS-EMP I-O tables appear to have different levels of aggregation, with the 1977-1985 series covering 185 sectors and the coverage of the 1960-85 series left unspecified.

Next, Jorgenson et al. (2013) offers a brief methodological description of how they consolidated the detailed sectoral coverage into the BLS-EMP I-O data into the 35 sectors used in the Jorgenson data, based on the detailed work Jorgenson and his colleagues

48 Id. at 452.

49 Id. at 451, FN 7 (indicating the data are available from “the Office of Occupational Statistics and Employment Projections website at www.bls.gov/emp”). However, as of July 24, 2014, this website no longer has any historical I-O data available.

50 BLS, supra note 15.

51 Jorgenson et al. (2013), supra note 6 at 453-457 (Table B.3) (showing the mapping between Jorgenson’s 35 sectors and the 192 sectors identified by the legacy BLS-EMP I-O data tables).

52 Id. at 452. Jorgenson et al. (2013) suggest that the methodology for converting detailed sectoral coverage to their aggregate 35 categories is the same, whether the number of industries is 192, 185, or something else. Id.
have published in this field over the years.\footnote{Id. at 452, 458.} Once these BLS-EMP I-O data series have been converted to the 35-sector approach in the Jorgenson data, they are merged into an integrated series covering the period 1960-2000.\footnote{Id. It appears that Jorgenson et al. (2013) took the 1983-2000 series as a starting point. Next, they took the data for 1983 as an “initial guess” for the 1982 I-O tables, which was then used to begin the iterative proportional fitting method developed in Jorgenson et al. (1987), supra note 5. This process is repeated to generate the remaining years back to 1977. Jorgenson et al. (2013), supra note 6 at 452. It is then applied to generate coherent data back to 1960. Id.} The last step is to integrate the data for the period 2000-2005, a process that turns on the NAICS-SIC transition, as described below.

7. NAICS-SIC Transition

In the late 1990s, the Office of Management and Budget led a collaborative effort to develop the North American Industry Classification Scheme (“NAICS”), which replaced the older Standard Industrial Classification (“SIC”) system. Most federal agencies—including the Census Bureau, BEA, and BLS—have adopted the newer NAICS structure. The transition has important implications for historical analysis because the definitions of key industries changed.\footnote{U.S. Census Bureau, 1997 Economic Census: Bridge Between NAICS and SIC, available at 
https://www.census.gov/epcd/ec97brdg/.} As a result, one cannot simply identify an industry and presume that economic data referring to that industry perfectly correspond on both sides of the NAICS-SIC transition; more analysis is required to confirm how significant the statistical discrepancy is for the few years where primary data were collected in both systems.

As a reminder, the Jorgenson data are classified using definitions that “roughly” match 2-digit SIC classifications, yet cover a period that includes years in which primary government data was not collected in the SIC scheme.\footnote{35-Sector KLEM Instruction Sheet, supra note 8. In other words, each industry in the Jorgenson data series “roughly” corresponds to a 2-digit SIC definition. For example, this means that Jorgenson’s industry #20, “Primary Metal,” should roughly correspond to SIC 33 Primary Metals.} It is worth noting that neither BLS nor BEA have arrived at a permanent solution to fully harmonize the two accounting systems over time, with only a one-time bridge available from the Census Bureau.\footnote{U.S. Census Bureau, supra note 55.} Switching completely to NAICS is no panacea, either. While some agencies have put
significant effort into modeling how the original SIC data would have been reported in NAICS terms—an estimation, not new primary data—these NAICS data go back only as far as 1993 for BLS, and only as far as 1998 for BEA. Merely relying on SIC data is insufficient as well. BLS no longer publishes any annual I-O SIC data, and BEA discontinued its annual I-O SIC data after the transition to NAICS; BEA SIC data is available only from 1977 to 1997.

The Jorgenson data address these issues in a way that reflects the discontinuity of national accounting systems. Recall that the Jorgenson data are presented in SIC terms; using a combination of primary and secondary data, Jorgenson and his colleagues are able to estimate data for the period 1960-2000. In order to extend this series to 2005, additional assumptions are required. Jorgenson et al. (2013) begin with annual NAICS-based I-O tables published by BLS-EMP. Next, they look at gross industry output data published by the BEA for the period 1998-2005 at the 485 NAICS industry level. Jorgenson et al. (2013) then convert the BEA data from 485 NAICS sectors to the 35 sectors used in the Jorgenson data. But rather than use the official Census NAICS-SIC bridge—which

61 See Section 6, supra.
62 Jorgenson et al. (2013), supra note 6 at 458. Note that Jorgenson et al. do not specify a particular source, referring only to a BLS-EMP (“series of IO tables for 1998-2004 based on 200 NAICS industries.”) Id. In contrast, the only I-O tables currently published by BLS cover the period 1993-2012 at the 195 NAICS industry level. BLS-EMP, supra note 58 (containing a data set published on Dec. 19, 2013). It is possible that BLS-EMP modified its scope and level of industry detail after publication of the Jorgenson et al. (2013) book, but without a specific citation to a data source, we cannot locate the primary data. Nevertheless, the conceptual approach taken by Jorgenson et al. (2013) is straightforward.
63 Id. Jorgenson et al. (2013) cite the BEA “Gross Output by Industry” data purportedly available at http://www.bea.gov/industry/gdpbyind_data.htm (last visited July 23, 2014), but none of the data series published there match the description given in the Jorgenson et al. (2013) text. BEA does provide Gross Output by Industry for the period 1997-2012, but at the 402 NAICS industry level. Given that this is merely an expanded coverage and a slightly reduced level of detail, it is likely related to, or at least conceptually equivalent to, the primary source Jorgenson et al. (2013) used.
maps the complex re-allocation of industries across the two accounting systems—Jorgenson et al. (2013) assigned each of the 485 sectors 100% to one of the aggregated 35 sectors in the published Jorgenson data.\(^{64}\) Thus, this imperfect (but tractable) assumption introduces an unspecified error into the final Jorgenson data set. Next, Jorgenson et al. (2013) calculate the growth rates over 2000-2005 from the BEA NAICS data and apply these to the 2000 SIC gross output data.\(^{65}\) Finally, they calculate the annual I-O data for 2000-2005 by extrapolation from the 2000 Make and Use I-O tables.\(^{66}\)

8. Comparing Sources for Gross Output Data

Briefly, we note that the choice of primary data sources has significant implications for the content of the final I-O table used to create the sector-specific price and consumption data in the final Jorgenson data set. One of Jorgenson’s later books reviews alternative options for constructing gross output data, which are used along with the I-O data to construct a complete map of inter-industry transactions, and hence, the final pattern of energy consumption in the national economy. As Table SI-1 shows, different BLS and BEA primary data sources vary widely in their assessment of the average annual growth rate of gross sector output for the five energy-producing sectors in the Jorgenson framework. Because I-O tables require an exponentially greater number of data points than do estimates of gross output, the magnitude of the effect of selecting different data sources on gross output suggests that even greater differences are likely present in the I-O table.\(^{67}\) This should decrease our confidence in the assertion that any particular I-O table used by Jorgenson perfectly allocates the mixture of five energy commodities across the econo-

\(^{64}\) Jorgenson et al. (2013), supra note 6 at 458.

\(^{65}\) Id.

\(^{66}\) Id. (“We use the Make matrix from 2000 to derive SIC commodity values from the industry outputs and use the SIC-based industry output and commodity output series as control totals to extrapolate the 2000 Use and Make tables to 2005.”)

\(^{67}\) For example, if an economy has \(n\) industries, a data series measuring gross output by industry contains \(n\) data points. Because each industry produces commodities that can be exchanged with all other industries, a simple I-O table for this economy would have \(n^2\) data points.

Each entry in the I-O table, \(X_{ij}\), represents the flow from industry \(i\) to industry \(j\). By definition, the sum of all transactions from industry \(i\) to all sectors in the economy is directly related to the gross output of industry \(i\). Thus, any error in the gross output of industry \(i\) could be magnified across the I-O table. As a result, one should have less confidence in any particular \(X_{ij}\) than in the corresponding estimate of gross output for industry \(i\).
my. Again, Jorgenson’s methods may be perfectly reasonable assumptions for the purposes of macroeconomic analysis, but our examination reveals an important source of uncertainty in the application of this approach to the study of energy efficiency policy or rebound effects.

9. **Correspondence Between Jorgenson and EIA Energy Price Data**

Fundamentally, the energy prices reported in the Jorgenson data for each sector are based on inter-industry consumption of the primary outputs from each of five energy-producing sectors (Coal mining, Crude oil and gas extraction, Petroleum refining, Electric utilities, and Gas utilities). As a threshold matter, we note that two of the energy-producing sectors produce multiple energy commodity outputs. This complication frustrates any comparison of the national accounting measurements with primary energy commodity price data. First, the crude oil and gas extraction sector produces both petroleum and natural gas, with a variety of qualities and types of hydrocarbon resources in each category. Second, petroleum refining takes crude oil and petroleum coke, transforming these primary resources into a range of refined products—from gasoline to jet fuel, diesel to asphalt. For both sectors, relating an estimate of a sector-wide output prices to a single energy commodity price index is not realistic. In addition, government energy data collection does not measure composite average prices for the output of these complex sectors, focusing instead on the individual energy commodities traded in energy markets.

On the other hand, the remaining three energy-producing sectors are more readily compared to individual energy commodities. We rely on data from the U.S. Energy Information Administration to assess these sectors. First, coal mining produces coal, the price of which can be tracked at the national level, despite regional variation in coal quality, the presence of harmful pollutants (e.g. mercury and sulfur), and transportation prices.68 Second, electric utilities are in the business of selling electricity to customers. As discussed in the main text, the rate varies widely by customer segment and geography—but again, the U.S. government also tracks national average prices.69 Third, gas utilities

---


are in the business of selling natural gas to customers. Here, the distinction between customer classes is equally significant, but the government does not estimate average natural gas prices. Instead, EIA tracks national average wellhead prices, changes to which provide a reasonable proxy for changes to the average price all consumers receive.

We compare the relationship between energy prices reported for three of the five energy-producing sectors in the Jorgenson data with government price data for the roughly equivalent energy commodities—coal mining and coal (Figure SI-3), electric utilities and retail electricity (Figure SI-4), and gas utilities and wellhead natural gas prices (Figure SI-5). As can be seen, the correspondence between the EIA and Jorgenson data is uneven and occasionally very weak. This may be due in part to the imperfect relationship between the primary industry output in Jorgensohn’s classification scheme and the energy commodity against which each industry’s price series is compared. In turn, this mismatch reflects the different accounting methodologies employed by Jorgenson (who focuses on national accounting industrial classifications) and primary energy data from EIA (which focuses on energy commodities, fuel types, and relevant customer classes, rather than composite industrial output indices). In the case of electric and gas utilities, the imperfect relationship likely reflects the fact that Jorgenson was unable to track these two industries directly, and instead inferred them in part by reference to an aggregated utility category.

While different accounting systems need not match one another perfectly at every level of detail, the fact that the two data series diverge significantly where there should be conceptual overlap raises additional questions about the appropriateness of Saunders’ use of the Jorgenson data to investigate the extent of the rebound effect—especially without any consideration of uncertainty or even a simple sensitivity analysis.

---

70 *Id.* at Table 9.10.

71 Wellhead prices offer an upstream proxy that is not directly comparable to the price paid by end-use consumers. We selected this series because while EIA traces some end-use price series, the government does not aggregate them into a composite end-use price estimate. Different customer classes face significantly different prices. We chose a simpler proxy (wellhead prices) in order to avoid making additional assumptions about the size of each-use consumption segment, assumptions that would have been necessary to generate a reliable downstream price series. Additional analysis is needed to fully compare differences between the Jorgenson price and official data across different end-use customer segments.
10. Data Quality Conclusions

Although we were unable to locate the primary data sources used to generate the Jorgenson data, the available documentation clarifies their conceptual origins. The primary sources are four benchmark I-O surveys from BEA (1977, 1982, 1987, and 1992). These surveys are supplemented by three secondary sources: annual I-O series from BLS-EMP, which interpolated between benchmark BEA survey years to cover a broader period (1960-1985, 1977-1995, and 1983-2000, respectively). The level of sectoral detail varied across primary and secondary sources, but was in each case (1) more detailed than the 35 sectors used in the Jorgenson data, and (2) based on the SIC national accounting system. As a result, Jorgenson needed to covert these data into his 35-sector scheme before combining them to create a coherent series that covers the period 1960-2000. To obtain the last few years of data, Jorgenson necessarily confronted the change in national accounting systems from NAICS to SIC. He calculated growth rates from NAICS-based data and applied these factors to the final year of SIC gross output data; he and his colleagues then extrapolated the I-O matrix using a method described in his 1987 book. These final extrapolated years were added to the earlier series, resulting in the final Jorgenson data covering the period 1960-2005.

Despite the incomplete documentation of the ultimate data sources—the three BLS-EMP sources are never fully specified, and were therefore impossible for us to obtain—Jorgenson and colleagues provide a clear methodological structure. Estimates of gross output by industry are combined with annual I-O tables to produce a complete matrix of inter-industry transactions. Jorgenson then applies his theoretical production function to these data to produce his published data. Crucially, this production function assumes that every purchaser faces the same price for each commodity—in other words, that there is only one price for each commodity every year. Because the I-O tables indicate that each sector consumes different baskets of intermediate commodities, and because energy inputs are defined as an aggregate category of five different output sectors, each sector will be associated with a unique energy price under this formulation. Nevertheless, that differentiation masks a fundamental assumption—that there is one price index for coal, one price index for oil and gas (combined), and so forth. In reality, of course, the price of energy varies widely by location, customer type, and by contractual arrangement.
In addition, because the underlying data sources do not track electricity utilities and gas utilities separately, Jorgenson must make additional assumptions to separate the combined “utilities” category into these two components. Specifically, he assigns fractional shares of the aggregate category to electric and gas utility sub-categories on the basis of their historical relationship between 1977-1995—which he calculates as 90% to gas utilities, and 10% to electric utilities. Thus, flows of two of the five energy inputs in Jorgenson’s framework are necessarily inferred by assumption, not direct observation; and the methodology for inferring the disaggregation is never fully disclosed.

Finally, we note that for three of the five energy-producing sectors whose primary output corresponds to a conventional energy commodity, the prices reported in the Jorgenson data vary significantly from those tracked by the U.S. Energy Information Administration.
This table presents the average annual rate of growth of gross output for the five energy-producing sectors in the Jorgenson data, as determined by four different primary data sources Jorgenson considered adopting. Jorgenson ultimately chose the BLS-EMP series. As this table shows, however, the different primary data sources vary widely in their estimations of the average annual growth rate of gross output over the period 1987-2000 for the five energy sectors in Jorgenson’s framework. Note that the sector-specific energy prices in the Jorgenson data are based on calculations derived from I-O tables, which are significantly more complex than tables of gross output by sector. As a result, the choice of data series should have even greater impact on the composition of annual I-O tables. See the full table in Jorgenson et al. (2005) for a comparison of all 41 sectors in that study across BLS and BEA data sources.

Table SI-1: Comparison of average annual growth in gross output by sector and data source (% per year over 1987-2000).\(^72\)

<table>
<thead>
<tr>
<th>Industry</th>
<th>Data Source</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>BLS-EMP</td>
</tr>
<tr>
<td>Coal Mining</td>
<td>1.19</td>
</tr>
<tr>
<td>Oil &amp; Gas Mining</td>
<td>–0.19</td>
</tr>
<tr>
<td>Petroleum &amp; Coal Products</td>
<td>0.97</td>
</tr>
<tr>
<td>Electric Utilities</td>
<td>2.12</td>
</tr>
<tr>
<td>Gas Utilities</td>
<td>–1.52</td>
</tr>
</tbody>
</table>

\(^72\) Jorgenson et al. (2005), supra note 34 at 116-17 (see Table 4.4).
Figure SI-1: Fuel shares for the commercial sector (EIA data). This figure shows the share of natural gas and retail electricity in the final energy consumption mix for the commercial sector. Electricity’s share has continuously grown in importance in the commercial sector since World War II. In contrast, the relative share of natural gas grew even more quickly at first, peaking in the mid-1960s, and then declined more slowly thereafter. The grey area highlights the period from which Jorgenson apparently extrapolates the share of output from Electric Utilities and Gas Utilities from an aggregate category.


Note that retail electricity does not include generation, transmission, and distribution losses; it represents final demand as calculated in kWh, not primary energy demand.
Figure SI-2: Fuel shares for the industrial sector (EIA data).  
This figure shows the share of natural gas and retail electricity in the final energy consumption mix for the industrial sector. Electricity’s share grew slowly but steadily in the industrial sector since World War II, with growth slowing over the last several decades. In contrast, the relative share of natural gas has risen and fallen several times over the same stretch of time. The grey area highlights the period from which Jorgenson apparently extrapolates the share of output from Electric Utilities and Gas Utilities from an aggregate category.

\[74\]  
*Id.* at Table 2.1d.

Note that retail electricity does not include generation, transmission, and distribution losses; it represents final demand as calculated in kWh, not primary energy demand.

Note also that petroleum refining is a significant industrial energy user; this industry consumes mostly crude oil and process gases produced during refining, which helps explain the relatively modest total share of gas and electricity in this chart.
This figure compares Jorgenson’s data on the average sector output price for Coal mining (sector #3) (“pz3” in the data file) with EIA data on the average national price for coal. EIA prices are consistently higher for the first four decades in this period. For consistency with the published Jorgenson data, prices are shown using a nominal index, with the value for 1996 = 1.0 by definition.

U.S. Energy Information Administration, supra note 68 at Table 7.9.
Figure SI-4: Comparing the Jorgenson and EIA data for electricity. 

This figure compares Jorgenson’s data on the average sector output price for Electric utilities (sector #30) (“pz30” in the data file) with EIA data on the average national price for retail electricity. EIA prices are significantly higher through the mid-1970s, but are eclipsed by a rapid rise reported in the Jorgenson data that persists through the mid-1980s. The two series then converge somewhat for another decade or so, with Jorgenson’s reported data rising rapidly at the end of the period. For consistency with the published Jorgenson data, prices are shown using a nominal index, with the value for 1996 = 1.0 by definition.

---

76 U.S. Energy Information Administration, supra note 69 at Table 9.8.
Figure SI-5: Comparing the Jorgenson and EIA data for natural gas.\textsuperscript{77}

This figure compares Jorgenson’s data on the average sector output price for Gas utilities (sector #31) (“pz31” in the data file) with EIA data on the average national price for natural gas, as measured at the wellhead. For the first two decades, Jorgenson’s price data are significantly higher than the EIA data. In the 1980s through the mid-1990s, the two series converge and move in tandem. In the late 1990s, however, the Jorgenson data do not reflect the significant increase in natural gas prices observed in the United States at that time. For consistency with the published Jorgenson data, prices are shown using a nominal index, with the value for 1996 = 1.0 by definition.

\textsuperscript{77} Id. at Table 9.10.