Regional Research Institute
West Virginia University

Working Paper Series

Value Chains: Production Upstreamness and Downstreamness Revisited

Patricio Aroca and Randall Jackson

Working Paper Number 2018-01_R1

Website address: rri.wvu.edu

Value Chains: Production Upstreamness and Downstreamness Revisited

Patricio Aroca*[1] and Randall Jackson†[2]

[1]CEPR, Business School, Universidad Adolfo Ibáñez, CHILE
[2]Regional Research Institute, West Virginia University, USA

November 19, 2018

Abstract

Measures devised to quantify value chain position have been used increasingly in recent years. While the constructs underlying these measures are meaningful, this paper identifies an overlooked implementation problem. Proposed algorithms have been applied as though the underlying data represent flows. Implementation data are drawn from modern input-output accounting frameworks that recognize secondary production explicitly. Unadjusted Use matrices are not conventional flows matrices because they do not identify the industries from which commodities originate. We demonstrate logical inconsistencies that arise, provide correct flow matrix formulations for upstreamness and downstreamness measures, and present empirical comparisons of correct and incorrect formulations.

Keywords: Value Chains, Input-Output, Upstreamness, Downstreamness
JEL Classification: F10,F14,L16,L23,C67

*patricio.aroca@uai.cl.
†rwjackson@mail.wvu.edu

The characterization of global supply chains is a topic that has gained in visibility and importance in the literature, particularly during the last decade. Various authors have approached supply and value chains from perspectives that include business transaction optimization, economic development policy, and property rights. An important case in point is a paper by Antràs and Chor (2013) (AC) that addressed the problem of measuring the degree to which industries are upstream or downstream in the global value chain, including an algorithm that they developed and used for these purposes. Since the 2013 AC publication and its companion, Antràs et al. (2012) (ACFH), a growing number of publications have built upon the incorrect measurement formulation (examples include Giovannetti et al. (2015); Carvalho (2014); Engemann et al. (2014); Gallego et al. (2015); Amador (2016); Amador J and Steher (2015); Ju and Yu (2015); Kee and Tang (2016); Matous and Todo (2016); Nagengast and Stehrer (2016); Temurshoev and Oosterhaven (2014); Timmer et al. (2015, 2014); Zhu et al. (2015)). Unfortunately, their algorithmic implementation fails to recognize and account for the differences and distinctions between historical input-output (IO) accounting frameworks based on interindustry flows matrices, and modern supply–use, or commodity-by-industry accounting frameworks. Further, the explicit recognition of the realities of modern accounting frameworks makes explicit the opportunity and need for researchers to identify differences in goals and objectives that might lead to the selection of alternative flows matrix formulations for identifying *upstreamness* and *downstreamness*.

The purpose of this paper is to identify explicitly the conceptual differences between the implementation algorithm and the underlying accounting frameworks, to demonstrate that these distinctions can have substantial empirical manifestations, and to provide correctly formulated alternatives.

Further, we emphasize that despite the well-established and well-understood distinctions between interindustry accounts and commodity-by-industry accounts, those who develop IO-related or IO-based methods all too often begin their conceptual formulations with interindustry accounts instead of the supply-use framework that is the format in which IO data for most developed nations are published. Not only can this lead to theoretical, conceptual and empirical inconsistencies, but the additional information in the supply-use framework can sometimes provide excellent opportunities for problem solutions that are not available when we limit ourselves only to interindustry accounts.

The remainder of the paper is organized as follows. In the next section we describe in general terms the context for the paper by revisiting the origins and objectives of the organizational framework for modern IO accounts. The third section provides empirical foundations for treating secondary production consistently. Section 4 focuses on the mathematical formulation of the upstreamness measure, and identifies the interpretive inconsistencies that arise when the Use matrix is used as though it were a flows matrix. Section 5 presents the correct mathematical formulations, which are followed in Section 6 by the comparisons of empirical outcomes from the alternative implementation methods. Section 7 provides a summary and discussion.

1 Misinterpretation and Misuse of the Commodity-by-Industry Accounting Framework

Richard Stone received the 1984 Nobel prize for his contributions to the design of the modern commodity-by-industry accounting framework that was adopted by the United Nations (UN) a half century ago (United Nations, 1968). Stone (1961) laid the foundations for the data organization upon which the UN System of National Accounts (SNA) rests.[1] Despite the Nobel recognition and the long-standard commodity-by-industry framework for national accounting, the fundamentals of data structures in published national accounts often seem either to be ignored or poorly understood by many of today's economists. Traditional IO tables published prior to the adoption of the UN SNA framework were symmetrical and did not account for the significance of secondary production. Industries and their outputs were treated as though they were homogeneous. Interindustry accounts detailed interindustry transactions irrespective of the mix of products that each industry produced. The UN SNA distinguished between producing industries and the products (commodities) they produced, and promoted the nearly globally accepted approach of accounting separately and independently for industry purchasing patterns (Use matrix) and output distributions (Make or Supply matrix). This shift in organizing the underlying economic system data made it possible to allocate properly the costs of production, given one of two assumptions about secondary production technology. All of this was

[1] The UN SNA is not alone in adopting this framework. The OECD has adopted the Supply-Use framework (OECD (2018)), and the recently developed World Input-Ouput Database (Dietzenbacher et al. (2013)) organizes its data similarly.

developed in the context of the double entry accounting principles that assure the equality and consistency of the input and output balance relationships in the system of sales and purchases.

The papers by Antràs and his colleagues have introduced to the field an approach that, while correct in construct, overlooks the fundamental accounting relationships and identities of the current generation of IO accounting frameworks. In so doing, they have laid a potentially problematic foundation for a set of upstream and downstream production system linkage measures that is based on data organized under the Stone accounting system.

2 Commodity Versus Industry: A Necessary Distinction

The statement "$d_{ij}Y_j$, is precisely the value of commodity i used in j's production" (ACFH, page 414) refers to their formulation of the coefficient d_{ij} as an element of a Use table, and Y_j as industry j output. However, this definition is somewhat obscured by the fact that industries produce more than one commodity. Hence, d_{ij} would have a very different interpretation, and a different value, if it were drawn from a commodity-by-commodity IO table, in which the denominator would have the value of commodity rather than industry output. The choice of the denominator is critical to the interpretation and the value of d_{ij}, especially because the gross output of industry j can be very different from the gross output of commodity j. For example, the most upstream industry in their analysis, which uses 2002 U.S. data, is the the Petrochemical sector. Petrochemical *commodity* gross output is 22% larger than petrochemical *industry* gross output. This is because other industries also produce petrochemical commodities as a secondary product. The assumption that commodity and industry gross output are the same is invalid.

The definitions of measures in AC are equally imprecise. For example, their "first measure is the ratio of the aggregate direct use to the aggregate total use (DUse_TUse) of a particular industry i's goods, where the direct_use for a pair of industries (i, j) is the value of goods from industry i directly used by firms in industry j to produce goods for final use, while the total use for (i, j) is the value of goods from industry i used either directly or indirectly (via purchases from upstream industries) in producing industry *j*'s output for final use" (p. 2131). The *direct use* definition would have to be based on an industry-by-industry intermediate transaction matrix, while the *total use* would have to be based on an industry-by-industry Leontief inverse matrix post-multiplied by a diagonal matrix of the total output by industry. Nevertheless, they draw their values from a Use matrix to implement their measures. Neither reference to "the value of goods from industry *i*" are accurate when the unadjusted Use matrix is used as the data source.

To emphasize the importance of these distinctions, focus again on the Petrochemicals industry and commodity. Table 1 of Antràs et al. (2012), reproduced below as Table 1, identifies Petrochemicals (325110) as the industry with the highest upstreamness measure. However, the Make matrix that corresponds to the Use table used for that analysis indicates that only 36% of this industry's total gross output is its primary commodity – petrochemicals. The rest of the petrochemical industry's gross output is secondary production, i.e., production of other commodities. From the commodity output perspective, the petrochemical industry is the source of only 44% of petrochemicals commodity production. The rest of the petrochemicals commodity production comes from other industry sectors as secondary production.

Table 1: Least and Most Upstream Industries (Manuf.)

US IO2002 industry	Upstreamness
Automobile (336111)	1.000
Light truck and utility vehicle (336112)	1.001
Nonupholstered wood furniture (337112)	1.005
Upholstered household furniture (337121)	1.007
Footwear (316200)	1.007
Alumina refining (33131A)	3.814
Other basic organic chemical (325190)	3.853
Secondary smelting of aluminum (331314)	4.064
Primary smelting of copper (331411)	4.355
Petrochemical (325110)	4.651

Source: Authors' calculations, replicating Antràs et al. (2012).

Secondary production is precisely the reason why the commodity-by-industry framework for IO accounting was developed. Ignoring these definitional differences can result in substantial errors, misattributions, and misinterpretations. Therefore, we clarify and make explicit the relevant definitions, and describe the ways the variables should be used in studying a supply or value chain. The choice of model structure will depend on whether the goal of the analysis is to identify steps in the production of specific commodities, or inter*industry* linkages, i.e., whether the focus is on petrochemicals as a commodity, or on the linkages of the petrochemical industry with other industries. If it is the former, then a commodity-by-commodity IO specification would be appropriate, while if it is the latter, one would presumably use an industry-by-industry IO table.[2] We review below the ACFH presentation so that we can assess whether the definition that is used in the paper relates to any of these alternatives, and if not, whether it is appropriate for their demonstration analysis.

3 THE UPSTREAMNESS MEASURE

To develop their upstreamness measure, ACFH "begin by considering an N-industry closed economy with no inventories. For each industry $i \in 1, 2, \cdots, N$, the value of gross output (Y_i) equals the sum of its use as a final good (F_i) and its use as an intermediate input to other industries (Z_i)" (p. 412)

$$Y_i = F_i + Z_i = F_i + \sum_{j=1}^{N} d_{ij} Y_j \tag{1}$$

They define d_{ij} as "the dollar amount of sector i's output needed to produce one dollar's worth of industry j's output." This balance equation is correctly presented, provided that Y, F, and Z are industry rather than commodity sectors, and that the term "sector i" in the definition of d_{ij} refers to *industry* sector i.

They next derive equation (2), which is the Leontief inverse matrix expressed as an infinite series of terms that can be reduced to $Y = (I - D)^{-1}F$, where $(I - D)^{-1}$ is the traditional Leontief inverse. Based on these relationships and definitions, they then present their upstreamness measure, and

[2] A post-processing option for some policy purposes might be to conduct the analysis in a commodity-by-commodity framework, and then convert to industry space using the standardized Make table to transform from commodity output to industry output, as described in Section 4, below.

describe it as "the (weighted) average position of an industry's output in the value chain, by multiplying each of the terms in" (p. 413) the infinite series by their distance from final use plus one and dividing by the gross output of the industry. For industry i, this yields

$$U_{1i} = 1\frac{F_i}{Y_i} + 2\frac{\sum_{j=1}^{N} d_{ij}F_j}{Y_i} + 3\frac{\sum_{j=1}^{N}\sum_{k=1}^{N} d_{ik}d_{kj}F_j}{Y_i} + 4\frac{\sum_{j=1}^{N}\sum_{k=1}^{N}\sum_{l=1}^{N} d_{il}d_{lk}d_{kj}F_j}{Y_i} + \cdots \tag{2}$$

They go on to provide a computational reduced form for this measure, after establishing its equivalence to Fally's (2011) upstreamness measure, which has the following compact expression:

$$U_1 = (I - \Delta)^{-1} \cdot \mathbf{1} \tag{3}$$

"where Δ is the matrix with $\frac{d_{ij}Y_j}{Y_i}$ in entry (i, j) and $\mathbf{1}$ is a column vector of ones."

After providing two economic interpretations for these upstreamness measures – in which the emphasis appears to be linkages among industries and not products – they present an open-economy adjustment, justified by noting that the data used to construct their matrix of US IO coefficients "do not distinguish between flows of domestic goods and international exchanges" (p. 414). The result is an adjustment factor for the IO coefficients matrix that transforms its interpretation from a technical relationship to a trade relationship. The adjustment factor is the ratio of domestic output of industry i to domestic use (absorption) of industry i output. Their presentation to this point is correct.[3]

The conceptual inconsistency arises in the implementation of the upstreamness measure, which proceeds by replacing $d_{ij}Y_j$ in matrix Δ with U_{ij}, the ij^{th}cell from the Use matrix. Despite the recognition in ACFH footnote 3 that "the coefficient d_{ij} is computed as the total purchases by industry j of *industry* i's output," they draw coefficients from the Use table, which represents output of *commodity* i required to produce industry j's output, not *industry* i output, and the difference between industry and commodity output, as we have seen, can be substantial. More importantly, because

[3] Dietzenbacher and Romero (2007) developed and reported measures that are virtually identical to these measures.

commodities are produced by multiple industries, the Use matrix is not a flows matrix as were the historical interindustry transactions or IO coefficients matrices. While the columns of the Use matrix can be conceived of as destinations, the rows identify the commodities that are flowing but not their industries of origin.

The inconsistency can be clarified further by returning to the derivation of the reduced form upstreamness expression from the power series expansion, which was derived from the accounting identity $Y_i = F_i + Z_i = F_i + \sum_{j=1}^{N} d_{ij} Y_j$. Consider the $d_{il} d_{lk} d_{kj}$ term of ACFH equation (2), where these d coefficients are defined as ratios of industry input dollar per industry output dollar. If we assign values of .1, .2, and .3 to these coefficients, then every dollar of output from industry j will require \$0.3 of input k, which will create a requirement for $\$(.2)(.3) = \$0.06$ of input l, which will then require $\$(.1)(.2)(.3) = \$0.006$ of input i for its production. The numerators and denominators have the same dimension (industry \$), so the interpretation of the product is clear and consistent. If, however, the d_{ij} coefficients have industry denominators but *commodity* numerators, then the product now reflects commodity i required to produce industry l output times commodity l required to produce industry k output times commodity k required to produce industry j output. But because each of these industries produces secondary products, the one-to-one relationship is lost; the product of this multiplication makes sense dimensionally, and therefore has an unambigouously straightforward interpretation *if and only if* these industries produce only their own commodities, which would mean that industry and commodity output would have to be identical. This kind of system would be reflected in a Make table with nonzero elements only on the diagonal. Were these coefficients defined with commodity terms in both numerator and denominator, they would be interpreted as commodity i required in the production of commodity l times commodity l required in the production of commodity k times commodity k required to produce commodity j, and this would be dimensionally consistent.

The crux of the problem is that commodity required to satisfy industry demand results not only in the production of the industry's primary commodity, but also the production of secondary commodities, and this happens at every term in the power series expansion, resulting in the loss of ability to trace commodities unambiguously through the supply chain.

4 Upstreamness Reformulated

The values in the Use table are associated – behaviorally – with columns. They represent column industry requirements of row commodity inputs, so standardizing by commodity values is not particularly useful. This does not render the development of an ACFH upstreamness measure intractable, of course. The modern accounting system that is the commodity-by-industry framework was devised precisely to accommodate the need to work analytically with systems in which industries produce multiple commodities. Indeed, developing the requirements coefficients in either industry-by-industry or commodity-by-commodity format is possible using the same database that ACFH used to implement their measure. We provide below the fundamental accounting equations that support the construction of these requirements coefficients tables.

The commodity-by-industry framework is presented below in Figure 1. In conventional IO notation (Miller and Blair, 1985), the matrix partition $U = [u_{ij}]$ is the Use matrix, $V = [v_{ij}]$ is the Make matrix, e is commodity final demand expressed here as a single vector, q is commodity gross output, and g is industry gross output. Only in highly unusual cases will an industry produce no secondary commodities, so rarely will q_i and g_i be equal. The va term denotes value added.

The traditional industry output balance equation that ACFH write as $Y_i = F_i + Z_i$ actually has no simple and direct counterpart in the modern accounting framework (although one can be derived, it requires assumptions about secondary production technology and information contained in V).[4] We can, however, express a commodity output balance equation in this conventional notation, will be: $q_i = \sum_{j=1}^{N} u_{ij} + e_i$, and we can further define $d_{ij} = \frac{u_{ij}}{g_j}$ and substitute to obtain $q_i = \sum_{j=1}^{N} d_{ij} g_j + e_i$, maintaining the output balance.

In matrix notation, we have the following identities:

$$U\mathbf{i} + e \equiv q \tag{4}$$

$$V\mathbf{i} \equiv g \tag{5}$$

$$V'\mathbf{i} \equiv q \tag{6}$$

where $\mathbf{i}$ is a summing vector, and $'$ signifies the transpose operation. We define behavioral relationships as follows:

[4]The industry output balance equation is conventionally denoted X=Y+Z.

	Commodity	Industries	Final Demand	Totals
Commodity		U	e	q
Industries	V			g
Primary Inputs		va		
Totals	q'	g'		

Adapted from United Nations (1968)

Figure 1: The Commodity-Industry Framework

$$B = U\hat{g}^{-1} \tag{7}$$
$$U = B\hat{g} \tag{8}$$
$$D = V\hat{q}^{-1} \tag{9}$$
$$V = D\hat{q} \tag{10}$$

where ˆ indicates diagonalization. Equation 7 defines the production requirements of commodities per industry output dollar, and equation 9 is a statement of the industry-based technology assumption that commodities are produced by industries in fixed proportions.[5] Note that the effect of pre-multiplication of a commodity vector or matrix by D results in a transformation from commodity-space to industry-space, so $V\mathbf{i} = g = Dq$. This system allows us to formulate the following:

$$q = Bg + e \tag{11}$$
$$q = BDq + e \tag{12}$$
$$q = (I - BD)^{-1}e \tag{13}$$

The BD term is a commodity-by-commodity requirements matrix counterpart to the classical, column-standardized interindustry Leontief IO coefficients matrix. To convert to a Ghoshian (sales) matrix, we can express the

[5]The alternative is the commodity-based technology assumption, which while not used here, could be developed in parallel fashion.

coefficients in transactions as UD, then standardize the rows by commodity output, q, or $\hat{q}^{-1}UD$. The difference between the incorrect $\hat{q}^{-1}U$ formulation and the correct reformulation is quite clearly the D term, which is the essential mechanism that transforms the industry column dimension of the Use matrix to the commodity output space of vector q. The other difference is that ACFH convert commodity output to commodity product absorption by netting out trade and inventories, which can be implemented similarly in the new commodity-by-commodity requirements matrix reformulation by adjusting q for net trade and inventory adjustment before using it as a standardizing vector.[6]

5 Empirical Example

With the conceptual distinction established, we turn in this section to a demonstration of the empirical consequences for the computed measures reported in ACFH. Table 2 displays upstreamness scores using the ACFH method and data, and the corrected algorithm presented in this paper. The top seven rows of data report results for the commodity sectors whose scores rank among the top five using either method, and the last six rows correspond to the five lowest ranking commodity sectors using either method. Each row shows the scores for sectors that rank highest in the first data column, their ACFH ranks in the second data column, their scores from the reformulated algorithm in the third data column, and their respective ranks in the fourth.

Note first that there is a good deal of commonality. The ranks in the lowest ranking sectors are quite similar, which would be expected a) because of the correspondence between industries and commodities, especially for industries producing consumer products, although even here, Footwear Manufacturing shows an 8-point difference in ranks, and b) because sectors with the lowest scores are virtually never used as intermediate inputs and are therefore much less distinguishable in their scores. There is somewhat

[6] The $B\tilde{D}$ matrix, where $\tilde{D}$ is the Make matrix standardized by q adjusted for net trade and inventory, would be used in the computation of the counterpart, commodity-by-commodity downstreamness measure used in Antràs and Chor (2013). Although we have not addressed explicitly the derivation and use of the counterpart interindustry rather than inter-commodity measures, the development would follow a similar path but would be based on row- or column-standardization of the interindustry transactions matrix $\tilde{D}U$. This formulation based on the modification of the D matrix is introduced in Jackson (1998).

Table 2: Upstreamness measure comparisons

NAICS	Sector	ACFH Score	ACFH Rank	Corrected CxC Score	Corrected CxC Rank
325110	Petrochemical mfg	4.6511	1	4.1785	4
331411	Primary copper smelting and refining	4.3547	2	6.4031	1
331314	Secondary aluminum smelt & alloying	4.0637	3	3.9991	6
325190	***Other basic organic chemical mfg***	***3.8529***	***4***	***3.5391***	***15***
33131A	Alum. primary prodn and refining	3.8144	5	4.9632	3
331200	***Steel product from purchased steel***	***3.45***	***16***	***4.0065***	***5***
331419	***Prim. nonfer. metal smelt & refn***	***3.4186***	***18***	***5.6283***	***2***
336111	Automobile mfg	1.0003	279	1.0004	279
336112	Light truck and utility vehicle mfg	1.0005	278	1.0008	278
337122	Wood HH furniture mfg.	1.0052	277	1.0072	277
337121	Upholstered HH furniture mfg	1.0072	276	1.008	276
316200	***Footwear mfg***	***1.0073***	***275***	***1.0454***	***267***
336213	Motor home mfg	1.0123	274	1.0129	275

Source: Antràs et al. (2012) and authors‘ calculations. Largest ranking differences in bold.

greater discrepancy in the highest ranked sectors, however, where rank differences are as great as 16 among only the five most highly ranked upstreamness sectors. Indeed, the average difference in ranks is 7, due to the notable ranking shifts for Other Basic Organic Chemical Manufacturing, Steel Product Manufacturing from Purchased Steel, and Primary Smelting and Refining of Nonferrous Metals (except copper and aluminum).

The reformulation reveals substantial differences. Because of the multi-commodity reality of industry production, the second and fifth highest ranking upstreamness sectors using the correct method are not even in the top 15 sectors using the earlier formulation. Secondary production strongly influences the upstreamness measures, and cannot be ignored in implementation.

To demonstrate the inaccuracies introduced in counterpart downstreamness measures, we replicate AC's results for their DownMeasure (AC, page 2163) and provide in Table 3 corrected rankings for their highlighted industries. In Table 4 we provide the correctly ranked and highlighted industries that we obtain when using the correctly formulated commodity-by-commodity matrices. As with the upstreamness comparisons, there is some substantial agreement, but there also are some very substantial differences.

Table 3: DownMeasure Comparisons: AC Results

	Industry	AC DownMeasure	Rank	Corrected CxC Rank
Lowest 10 Values				
325110	Petrochemical mfg	0.2150	253	250
331411	Primary copper smelt & refining	0.2296	252	253
331314	Secondary aluminum smelt & alloy	0.2461	251	247
325190	Other basic organic chemical mfg	0.2595	250	241
33131A	Primary alumina refn and prodn	0.2622	249	251
325310	Fertilizer mfg	0.2658	248	248
335991	Carbon and graphite product mfg	0.2668	247	246
325181	***Alkalies and chlorine mfg***	***0.2769***	***246***	***224***
331420	Copper rolling, drawing, etc.	0.2769	245	244
325211	***Plastics material and resin mfg***	***0.2800***	***244***	***219***
Highest 10 values				
339930	***Doll, toy, and game mfg***	***0.9705***	***10***	***20***
311111	Dog and cat food mfg	0.9717	9	7
337910	Mattress mfg	0.9720	8	8
315230	***Women's and girls' apparel mfg***	***0.9762***	***7***	***17***
321991	Manufactured home mfg	0.9810	6	6
336212	Truck trailer mfg	0.9837	5	5
336213	Motor home mfg	0.9879	4	4
316200	Footwear mfg	0.9927	3	3
337121	Upholstered HH furniture mfg.	0.9928	2	2
336111	Automobile mfg	0.9997	1	1

Source: Antràs and Chor (2013) from authors' calculations. Largest rank differences in bold.

Table 4: DownMeasure Comparisons: Corrected CxC Results

	Industry	Corrected CxC DownMeasure	Corrected CxC Rank	AC Rank
Lowest 10 Values				
331411	Primary copper smelt & refining	0.1597	253	252
331419	***Primary nonferrous smelt &refn***	***0.1804***	***252***	***236***
33131A	Primary alumina refn and prodn	0.2049	251	249
325110	Petrochemical mfg	0.2467	250	253
331200	***Steel product mfg from purch steel***	***0.2560***	***249***	***238***
325310	Fertilizer manufacturing	0.2698	248	248
331314	Secondary aluminum smelt & alloy	0.2703	247	251
335991	Carbon and graphite product mfg	0.2735	246	247
333612	***Industrial drive and gear mfg***	***0.2743***	***245***	***205***
331420	Copper rolling, drawing, etc.	0.2750	244	245
Highest 10 values				
311230	***Breakfast cereal manufacturing***	***0.9629***	***10***	***15***
336612	Boat building	0.9700	9	11
337910	Mattress manufacturing	0.9725	8	8
311111	Dog and cat food manufacturing	0.9786	7	9
321991	Manufactured home mfg	0.9818	6	6
336212	Truck trailer manufacturing	0.9855	5	5
336213	Motor home manufacturing	0.9874	4	4
316200	Footwear manufacturing	0.9917	3	3
337121	Upholstered household furniture mfg	0.9931	2	2
336111	Automobile manufacturing	0.9996	1	1

Source: Antràs and Chor (2013) and authors' calculations. Largest rank differences in bold.

6 Implications for Practical Application

The need for the correction in formulation arises from the existence of secondary commodity production by industries. Hence, for an economy whose supply table is strongly diagonal – one with very little secondary production, empirical results might well differ little as a result of our correction, but the differences will become more substantial as the ratio of off-diagonal supply-table elements to diagonal elements increases. The degree of difference for a given set of accounts is an open empirical question, in that primary and secondary production structures vary geographically.

Because these metrics will most often be used in practice to identify and prioritize industries for further or attention, higher ranked industries will be of most interest and greatest value to anyone carrying out this kind of analysis. Therefore, the rank order correlations over the entire vector of ranks are not as relevant to the analyst as is the ability to correctly identify the top ranked industries. Below we present two additional ways of assessing the impacts correct formulation that underscore its importance in practical application.

First, we compute and display graphically the difference in ranks over the entire distribution of industries for which the measure has been calculated. Although correlations between the entire corrected and uncorrected upstreamness or downstreamness vectors can be quite high, the differences in ranks can be quite substantial. Figure 2 presents a plot of the simple differences in ranks for vectors containing corrected and uncorrected values for the upstreamness measure again derived from the same 2002 U.S. data used in ACFH, with sectors ordered according to their original industry classification scheme sequence.

To illustrate the importance of using the corrected formulation, then, we set up the following analysis, again using the same data. We first rank order sectors using the corrected measure values. We assign a value of 1 for$n = 1$ (where rank order calculations cannot be computed), then, as $n$increases, we select the top n corrected-ranked sectors from the incorrect vector, generate ranks for values within that set of n sectors, and then compute the Spearman rank correlations between vectors of increasing size. The result is a set of rank correlations for sets of n-dimension vectors. The correlations are actually best-case comparisons, because in most sets, uncorrected sector rank values from the entire set of industries will often exceed the value of n, but for the correlations to be valid, their ranks are indexed relative to the n values in

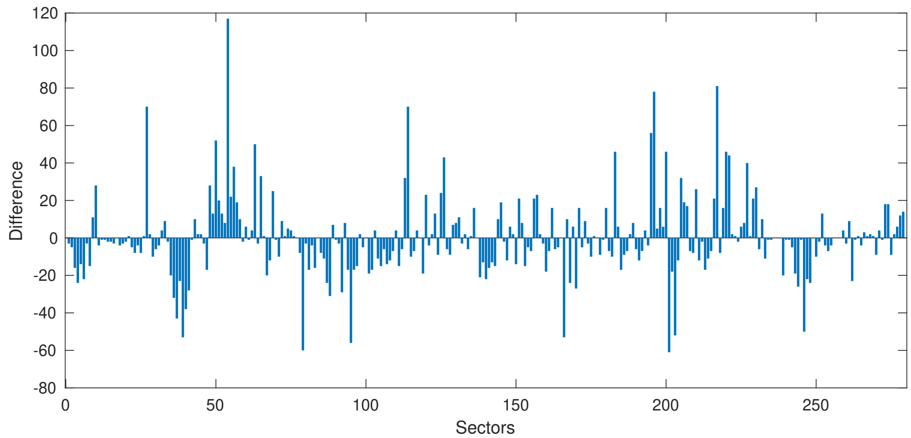

Figure 2: Upstreamness: Difference in Ranks

each vector in the comparison set. The result of that exercise for $n = 1$, ..., 50 is shown below in Figure 3, which reveals that at $n = 5$, there is zero correlation; for the top 10 ranked correctly calculated values, the correlation rises to 0.45, and by $n = 50$, the correlation is still only 0.58. While there is no clear way of assessing the these results comparisons statistically, the two ranks vectors are clearly not highly correlated enough to suggest that there is no substantial difference in the qualitative nature of the results, and certainly no support for simply ignoring the effect that the correction has on outcomes.

Irrespective of the empirical implications, of course, a published use table in isolation provides only a partial description of an IO system. This alone is reason enough to base empirical analyses on the correct formulations. The correction is straightforward and the data are virtually always published in tandem, so there is little reason not to use the correct formulation.

7 Summary and Discussion

In this paper, we identify an important inconsistency in the formulation and implementation of upstreamness and downstreamness measures developed and presented in AC and ACFH. The lack of correspondence between construct and data has carried through to the subsequent publications iden-

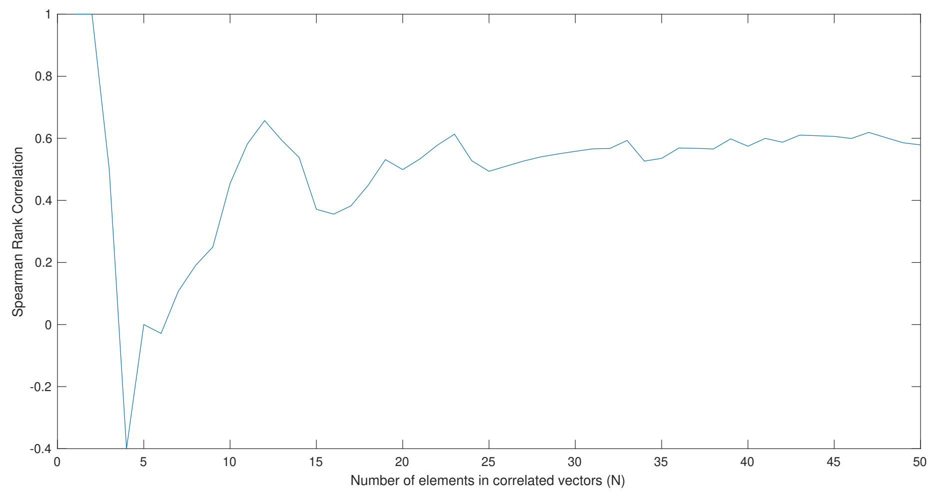

Figure 3: Upstreamness: Spearman Rank Comparison for Top n Ranks

tified in the introduction, necessarily generating inaccurate results. The problem arises because important differences between historical interindustry and modern commodity-by-industry IO accounting frameworks are overlooked. We develop and implement a correctly formulated alternative based on the commodity-by-industry framework that meets the objectives of the upstreamness and downstreamness measures. Our empirical demonstration makes clear the need to use the correct formulation when studying production linkages.

The options available to those studying upstream and downstream linkages in the context of supply and value chains are also worth considering, because each of these options produces a different kind of information. First, one can formulate these measures to study either commodity or industry chains. Commodity chain analyses will reveal information about the production of selected products, while industry supply chains can reveal information about the linkages among activities, specifically industries that are engaged in the production of one or more products. Industry and commodity linkages are surely different, and depending on the goal of the analysis, one or the other classification may be preferred. Second, the coefficients that define the interactions among commodities or industries can be purged of trade as in the procedures discussed here, and this provides a focus on the within-region (in this case, nation) production chains. However, the coefficients can also be based on technical requirements irrespective of origin, and this can provide information that can be useful in assessing the production structure of an economy relative to potential development alternatives.

Rather than making the accounting framework more complicated, the format of modern systems of national accounts actually enriches the possibilities for meaningful analysis, and facilitates analyses of products (commodities) and activities (industries) in economic systems. Time and effort spent deepening awareness and understanding of the underlying accounting conventions and embedded relationships can yield substantial research and application dividends.

References

Amador, J. (2016). Global value chains: A survey of drivers and measures. *Journal of Economic Surveys*, 30(2):278–301.

Amador J, Cappariello, R. and Steher, R. (2015). Global value chains: A view from the euro area. *Asian Economic Journal*, 29(2):21 pages.

Antràs, P. and Chor, D. (2013). Organizing the global value chain. *Econometrica*, 81(6):2127–2204.

Antràs, P., Chor, D., Fally, T., and Hillberry, R. (2012). Measuring the upstreamness of production and trade flows. *The American Economic Review: Papers and Proceedings*, 102(3):412–416.

Carvalho, V. M. (2014). From micro to macro via production networks. *Journal of Economic Perspectives*, 28(4):23–48.

Dietzenbacher, E., Los, B., Stehrer, R., Timmer, M., and de Vries, G. (2013). The construction of world input-output tables in the WIOD project. *Economic Systems Research*, 25(1):71–98.

Dietzenbacher, E. and Romero, I. (2007). Production chains in an interregional framework: Identification by means of average propagation lengths. *International Regional Science Review*, 30(4):362–383.

Engemann, M., Eck, K., and Schnitzer, M. (2014). Trade credits and bank credits in international trade: Substitutes or complements? *The World Economy*, 37(11):1507–1540.

Gallego, N., Llano, C., De La Mata, T., and DÍaz-Lanchas, J. (2015). Intranational home bias in the presence of wholesalers, hub-spoke structures and multimodal transport deliveries. *Spatial Economic Analysis*, 10(3):369–399.

Giovannetti, G., Marvasi, E., and Sanfilippo, M. (2015). Supply chains and the internationalization of small firms. *Small Business Economics*, 44(4):845–865.

Jackson, R. (1998). Regionalizing national commodity-by-industry accounts. *Economic Systems Research*, 10(3):223–238.

Ju, J. and Yu, X. (2015). Productivity, profitability, production and export structures along the value chain in china. *Journal of Comparative Economics*, 43(1):33–54.

Kee, H. L. and Tang, H. (2016). Domestic value added in exports: Theory and firm evidence from china. *American Economic Review*, 106(6):1402–1436.

Matous, P. and Todo, Y. (2016). Energy and resilience: The effects of endogenous interdependencies on trade network formation across space among major japanese firms. *Network Science*, 4(02):141–163.

Miller, R. E. and Blair, P. D. (1985). *Input-Output Analysis: Foundations and Extensions*. Prentice-Hall, Englewood Cliffs, New Jersey.

Nagengast, A. J. and Stehrer, R. (2016). Accounting for the differences between gross and value added trade balances. *The World Economy*, 39(9):1276–1306.

OECD (2018). Supply and use indicators. https://www.oecd-ilibrary.org/content/data/f5ff195e-en.

Stone, R. (1961). Input–Output and National Accounts (Paris, Organization for European Economic Cooperation).

Temurshoev, U. and Oosterhaven, J. (2014). Analytical and empirical comparison of policy-relevant key sector measures. *Spatial Economic Analysis*, 9(3):284–308.

Timmer, M. P., Dietzenbacher, E., Los, B., Stehrer, R., and de Vries, G. J. (2015). An illustrated user guide to the world input-output database: the case of global automotive production. *Review of International Economics*, 23(3):575–605.

Timmer, M. P., Erumban, A. A., Los, B., Stehrer, R., and de Vries, G. J. (2014). Slicing up global value chains. *Journal of Economic Perspectives*, 28(2):99–118.

United Nations (1968). *A System of National Accounts*, volume 2 of *Series F*. United Nations, New York, 3 edition.

Zhu, Z., Puliga, M., Cerina, F., Chessa, A., and Riccaboni, M. (2015). Global value trees. *PLoS One*, 10(5):e0126699.