NBER WORKING PAPER SERIES

# MEASURING THE UPSTREAMNESS OF PRODUCTION AND TRADE FLOWS

Pol Antràs
Davin Chor
Thibault Fally
Russell Hillberry

Working Paper 17819
http://www.nber.org/papers/w17819

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
February 2012

This paper was written for the invited session on “Global Production Chains”at the 2012 AEA meetings in Chicago. A shorter version of this paper is forthcoming in the American Economic Review Papers and Proceedings. We thank Jon Vogel for an insightful discussion, and David Atkin, Dave Donaldson and Marc Melitz for helpful comments. The data and programs associated with this paper are available from the authors' websites. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Measuring the Upstreamness of Production and Trade Flows
Pol Antràs, Davin Chor, Thibault Fally, and Russell Hillberry
NBER Working Paper No. 17819
February 2012
JEL No. F10,F14,L16,L23,O14

**ABSTRACT**

We propose two distinct approaches to the measurement of industry upstreamness (or average distance from final use) and show that they yield an equivalent measure. Furthermore, we provide two additional interpretations of this measure, one of them related to the concept of forward linkages in Input-Output analysis. On the empirical side, we construct this measure for 426 industries using the 2002 US Input-Output Tables. We also verify the stability of upstreamness across countries in the OECD STAN database, albeit with a more aggregated industry classification. Finally, we present an application that explores the determinants of the average upstreamness of exports at the country level using trade flows for 2002.

Pol Antràs
Department of Economics
Harvard University
1805 Cambridge Street
Littauer Center 207
Cambridge, MA 02138
and NBER
pantras@fas.harvard.edu

Davin Chor
Singapore Management University
School of Economics
90 Stamford Rd
Singapore 178903
Singapore
davinchor@smu.edu.sg

Thibault Fally
Department of Economics
University of Colorado at Boulder
256 UCB
Boulder, Colorado 80309-0256
USA
Thibault.Fally@colorado.edu

Russell Hillberry
Department of Economics
University of Melbourne
3010, VIC, Australia
Phone: +61 3 8344 5354
rhhi@unimelb.edu.au

# 1 Introduction

The fragmentation of production across national boundaries has been a distinctive feature of the world economy in recent decades. Production now often entails the sourcing of inputs and components from multiple suppliers based in several countries. These trends are likely to leave their imprint on international trade patterns: For example, are countries now specializing in particular stages of global production processes, or to borrow from Krugman (1995), specific slices of the value chain?

Addressing this question requires first and foremost an industry measure of relative production line position. In this article, we present two approaches to building a measure of industry "upstreamness" (or average distance from final use). The two approaches are motivated in distinct ways, but we prove that they yield an equivalent measure. Furthermore, we provide two additional economic interpretations of this measure, one of them closely related to the concept of forward linkages in Input-Output (I-O) analysis.

On the empirical side, we construct this measure using the 2002 US I-O Tables as a benchmark. The high level of disaggregation in the US Tables allows us to calculate upstreamness for a total of 426 industries. We separately construct our measure using the I-O Tables for selected OECD member countries from the STAN Database, in order to verify that upstreamness is a stable attribute of industries across different countries (with some caveats, see details in Section 4).

Finally, we present an application of our measure, by characterizing the average upstreamness of exports at the country level using trade flows in the year 2002. We also report here some regression-based findings which describe how upstreamness affects the cross-country, cross-industry pattern of trade. Our initial exploration indicates that stronger country institutions pertaining to the rule of law and financial development are correlated with a propensity to export in relatively more downstream industries. Our results also suggest a role of relative factor endowments in shaping the degree to which a country's exports appear to concentrate in relatively upstream versus downstream industries.

# 2 Three Measures of Upstreamness

## 2.1 Closed-Economy Benchmark

To build intuition, we begin by considering an $N$-industry closed economy with no inventories. In such an economy, for each industry $i \in \{1, 2, ..., N\}$, 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 ($Z_i$)

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

where, in the last summation, $d_{ij}$ is the dollar amount of sector $i$'s output needed to produce one dollar worth of industry $j$'s output. Iterating this identity, we can express industry $i$'s output as an infinite sequence of terms which reflect the use of this industry's output at different positions in the value chain, starting with final use

$$Y_i = F_i + \sum_{j=1}^{N} d_{ij} F_j + \sum_{j=1}^{N} \sum_{k=1}^{N} d_{ik} d_{kj} F_j + \sum_{j=1}^{N} \sum_{k=1}^{N} \sum_{l=1}^{N} d_{il} d_{lk} d_{kj} F_j + \ldots \quad (2)$$

Building on this identity, Antràs and Chor (2011) suggest computing the (weighted) average position of an industry's output in the value chain, by multiplying each of the terms in (2) by their distance from final use plus one and dividing by $Y_i$, or

$$\begin{aligned} U_{1i} &= 1 \times \frac{F_i}{Y_i} + 2 \times \frac{\sum_{j=1}^{N} d_{ij} F_j}{Y_i} + 3 \times \frac{\sum_{j=1}^{N} \sum_{k=1}^{N} d_{ik} d_{kj} F_j}{Y_i} \\ &\quad + 4 \times \frac{\sum_{j=1}^{N} \sum_{k=1}^{N} \sum_{l=1}^{N} d_{il} d_{lk} d_{kj} F_j}{Y_i} + \ldots \end{aligned} \quad (3)$$

It is clear that $U_{1i} \geq 1$ and that larger values are associated with relatively higher levels of upstreamness of industry $i$'s output. Although computing (3) might appear to require computing an infinite power series, provided that $\sum_{i=1}^{N} d_{ij} < 1$ for all $j$ (a natural assumption), the numerator of the above measure equals the $i$-th element of the $N \times 1$ matrix $[I - D]^{-2} F$, where $D$ is an $N \times N$ matrix whose $(i, j)$-th element is $d_{ij}$ and $F$ is a column matrix with $F_i$ in row $i$.[1] Using the fact that $Y = [I - D]^{-1} F$, which is easily verified from (1), the numerator can also be shown to equal the $i$-th element of the $N \times 1$ matrix $[I - D]^{-1} Y$, where $Y$ is a column matrix with $Y_i$ in row $i$.

Fally (2011) instead proposes a measure of upstreamness (or distance from final-good production) based on the notion that industries selling a disproportionate share of their output to relatively upstream industries should be relatively upstream themselves.[2] In particular, he posits the following linear system of equations that implicitly defines upstreamness $U_2$ for

[1] To be more specific, from the Perron-Frobenius theorems for non-negative matrices, the maximal eigenvalue of $D$ is bounded above by the largest column sum of $D$, which is less than one whenever $\sum_{i=1}^{N} d_{ij} < 1$ for all $j$. But if every eigenvalue of $D$ is less than one in absolute value, then $[I - D]^{-1}$ must exist.

[2] It should be noted that despite the order in which we introduce these measures, Fally (2011)'s measure chronologically precedes the one in Antràs and Chor (2011). Fally (2011) also proposes a measure of the number of stages embodied in an industry's output.

each industry $i$

$$U_{2i} = 1 + \sum_{j=1}^{N} \frac{d_{ij} Y_j}{Y_i} U_{2j}, \tag{4}$$

where note that $d_{ij}Y_j/Y_i$ is the share of sector $i$'s total output that is purchased by industry $j$. Again it is clear that $U_{2i} \geq 1$, and using matrix algebra, we can express this measure compactly as $U_2 = [I - \Delta]^{-1} \mathbf{1}$, where $\Delta$ is the matrix with $d_{ij}Y_j/Y_i$ in entry $(i, j)$ and $\mathbf{1}$ is a column vector of ones.

These two measures of upstreamness might appear distinct, but simple manipulations demonstrate that they are in fact equivalent (see the Appendix), which leads us to

**Proposition 1** $U_{1i} = U_{2i} = U_i$ *for all* $i \in \{1, 2, ..., N\}$.

A limitation of these two measures is that they impose an *ad hoc* cardinality in the sense that the distance between any two stages of production is set to one. In the Appendix we show, however, that these measures can in fact be given two precise economic interpretations. First, holding constant the final-use vector $F$ and the off-diagonal elements of the matrix $D$, we show that

$$U_i = \frac{1}{Y_i} \sum_{j=1}^{N} \frac{\partial Y_i}{\partial d_{jj}}, \tag{5}$$

so $U_i$ equals the semi-elasticity of an industry's output to a uniform change in input-output linkages *within* industries. Intuitively, when the extent to which industries' reliance on inputs from their own sector increases, this will tend to increase output in all industries, but one would expect the effect to be disproportionately larger in upstream industries via a multiplier effect.

Before providing a second economic interpretation of the upstreamness measure $U_i$, it is useful to note that in a closed economy with no inventories, the dollar value of gross output $(Y_i)$ in each industry $i \in \{1, 2, ..., N\}$ will also equal the sum of this industry's value added (or cost of primary factors, including profits), $V_i$, and its purchases of intermediate inputs from other industries, or using our notation above,

$$Y_i = V_i + \sum_{j=1}^{N} d_{ji} Y_i. \tag{6}$$

In the Appendix, we show then that, holding constant the allocation matrix $\Delta$, we have that

$$U_i = \sum_{j=1}^{N} \frac{\partial Y_j}{\partial V_i}. \tag{7}$$

Thus, $U_i$ also turns out to equal the dollar amount by which output of *all* sectors increases following a one dollar increase in value added in sector $i$. Readers familiar with Input-Output analysis will recognize that equation (7) relates the measure of upstreamness $U_i$ with a standard measure of cost-push effects or total forward linkages in supply-side Input-Output models (see Miller and Blair, 2009, particularly Chapter 12, for a discussion).[3] Again, it is intuitive that upstream industries will tend to generate more forward linkages in an economy, and that cost shocks in those industries will also tend to have a particularly magnified effect on prices (and on nominal output). What is more surprising, in our view, is that the measure of upstreamness independently derived by Antràs and Chor (2011) and Fally (2011) turns out to exactly equal this widely-known measure of forward linkages.[4]

## 2.2 Open-Economy Adjustment

So far we have assumed that the economy is closed to international trade. Since one of our main goals is to measure the level of upstreamness of a country's exports, it is important to extend the measurement of upstreamness to an open-economy environment. Incorporating this, the identity in (1) is now modified to

$$Y_i = F_i + \sum_{j=1}^{N} d_{ij} Y_j + X_i - M_i,$$

where $X_i$ and $M_i$ denote exports and imports of sector $i$ output.

It might appear that as long as net exports $X_i - M_i$ are not more or less upstream than domestic production, allowing for international trade flows would have no bearing on the measures of upstreamness discussed above. Nevertheless, it is important to emphasize that the interindustry commodity flow data used to construct the matrix of US input-output coefficients $D$ do not distinguish between flows of domestic goods and international exchanges.[5]

[3] Starting with Ghosh (1958), supply-side I-O analysis takes the allocation matrix $\Delta$ as the primitive of these models. In this literature, a standard measure of total forward linkages in a given industry is the corresponding row sum of the so-called Ghosh inverse matrix $[I - \Delta]^{-1}$. Remember, however, that Fally's upstreamness measure satisfies $U_2 = [I - \Delta]^{-1} \mathbf{1}$, which is precisely the vector of these row sums.

[4] Supply-side I-O models have been controversial in the literature because they appear to be based on assumptions inconsistent with Leontief's traditional demand-side models. As an example, our derivation of (7) holds the elements of the matrix $\Delta$ constant, but when output levels are endogenous, this can only occur if the elements of the matrix $D$ (the primitives of demand-side I-O models) adjust to "exogenous" changes in value added or cost of primary factors. Dietzenbacher (1997) proposed a price interpretation of supply-side models that reconciled these two approaches and also illustrated the links between forward linkages and cost-push effects.

[5] In other words, the coefficient $d_{ij}$ is computed as the total purchases by industry $j$ of industry $i$'s output, regardless of whether those purchases are domestic or involve imports. See Horowitz and Planting (2009) for more discussion, specifically the description of the Import Matrix in the I-O Tables. The OECD STAN data

Hence, although the share of a country's gross output in industry $i$ that is used as intermediate inputs in industry $j$ (at home or abroad) is given by the ratio

$$\delta_{ij} = \frac{d_{ij} Y_j + X_{ij} - M_{ij}}{Y_i}, \tag{8}$$

in practice we lack information on international interindustry flows $X_{ij}$ and $M_{ij}$.

It seems sensible, however, to assume

**Assumption 1:** $\delta_{ij} = X_{ij}/X_i = M_{ij}/M_i$.

In words, Assumption 1 imposes that the share of a country's industry $i$ output used in industry $j$ (at home or abroad), i.e., $\delta_{ij}$ in (8), is identical to the share of industry $i$'s exports (imports) that are used by industry $j$ producers. With this assumption, one can easily verify that our two measures of upstreamness in (3) and (4) continue to coincide after replacing $d_{ij}$ with

$$\hat{d}_{ij} = d_{ij} \frac{Y_i}{Y_i - X_i + M_i}. \tag{9}$$

Incidentally, the denominator in (9) is precisely the domestic absorption of industry $i$'s output. It is important to emphasize that although Assumption 1 imposes a certain structure on cross-country variation in production patterns, it is perfectly consistent with countries specializing in different segments of the value chain. We next illustrate this with a simple example that also highlights the importance of the adjustment in (9).

**Example.** Suppose that there are two industries, 1 and 2, and two countries, Home and Foreign. Industry 2 produces only intermediate inputs which are entirely sold to producers in sector 1, while sector 1 produces only final goods. Clearly, our closed-economy measure would suggest upstreamness values of 1 and 2 for industries 1 and 2, respectively. Suppose, however, that Home exports part of its production of good 1 to final consumers in Foreign, while Foreign producers of good 2 sell part of their output to Home producers in sector 1. Hence, relative to Foreign, Home appears to specialize in the relatively downstream sector. It is straightforward to verify (see the Appendix) that our adjusted measure delivers the correct values of upstreamness in each industry and each country (that is, 1 and 2), while, without the adjustment, the measure of upstreamness in industry 2 would be biased upwards at Home and biased downwards in Foreign, with the size of the bias increasing in the value of Foreign exports to Home. ∎

The above discussion abstracts from changes in inventories for ease of notation. A similar set of considerations is involved with inventories, as the input-output matrix $D$ does not

described below do have separate information available on import and domestic flows, but this information is often imputed under an assumption of proportional use of domestic and imported components.

separately identify inputs obtained from a draw-down of inventories as opposed to from fresh production. It is nevertheless straightforward to show that if we adopt a condition analogous to Assumption 1 in the treatment of inventories, then (9) is valid so long as $Y_i$ in the denominator is calculated subtracting the value of any net change in inventories of $i$ (see Appendix for details). This is in fact what we do in our empirical implementation below.

# 3 Upstreamness in US Production

We construct the above measure of industry upstreamness using the 2002 US benchmark Input-Output (I-O) Tables, as made available by the Bureau of Economic Analysis (BEA) on their website. A key advantage of the US data is that it reports information on production linkages between industries at a highly disaggregated level, namely at the level of six-digit I-O industry codes. There are altogether 426 industries in the I-O Tables, of which 279 are in manufacturing.

For our purposes, we use the detailed Supplementary Use Table after redefinitions. The $(i, j)$-th entry of this Use Table reports the value of inputs of commodity $i$ used in the production of industry $j$ in the US economy. An additional set of columns also records the value of commodity $i$ that enters into final uses, namely consumption, investment, net changes in inventories, and net exports.[6]

We construct the square matrix $\Delta$ with the open-economy adjustment in (9) as follows. The numerator of the $(i, j)$-th entry of $\Delta$, $d_{ij}Y_j$, is precisely the value of commodity $i$ used in $j$'s production; we therefore plug in the $(i, j)$-th entry from the Use Table for this numerator. The denominator $Y_i - X_i + M_i$ is in turn calculated as the sum of values in row $i$ of the Use Table, less that recorded under net exports and net changes in inventories. With this $\Delta$, the formula $[I - \Delta]^{-1} \mathbf{1}$ then delivers a column vector whose $i$-th entry is the upstreamness measure for industry $i$, as shown in Section 2.

The values we obtain reveal that industries vary considerably in terms of their average production line position. The measure of upstreamness ranges from a minimum of 1 (19 industries in which all output goes only to final uses) to a maximum of 4.65 (Petrochemicals). Its mean value across the 426 industries is 2.09, with a standard deviation of 0.85.[7] The average industry therefore enters into use in production processes roughly one stage before final consumption or investment. For illustrative purposes, Table 1 lists the 20 least and most upstream manufacturing industries. Of note, automobiles, furniture and footwear are

[6] The Use Table reports a further breakdown of the final use value of consumption and investment into private and government purchases. We will however not be using this breakdown in our analysis.

[7] These summary statistics are similar when restricting to manufacturing industries only.

Table 1: The Twenty Least and Most Upstream US Manufacturing Industries

| US IO2002 Industry | Upstreamness |
|---|---|
| Automobile (336111) | 1.0003 |
| Light truck and utility vehicle (336112) | 1.0005 |
| Nonupholstered wood household furniture (337112) | 1.0052 |
| Upholstered household furniture (337121) | 1.0072 |
| Footwear (316200) | 1.0073 |
| Motor home (336213) | 1.0123 |
| Truck trailer (336212) | 1.0165 |
| Manufactured home (mobile home) (321991) | 1.0194 |
| Women's and girls' cut and sew apparel (315230) | 1.0244 |
| Mattress (337910) | 1.0288 |
| Dog and cat food (311111) | 1.0291 |
| Doll, toy, and game (339930) | 1.0304 |
| Boat building (336612) | 1.0336 |
| Laboratory apparatus and furniture (339111) | 1.0383 |
| Blind and shade (337920) | 1.0413 |
| Electronic computer (334111) | 1.0428 |
| Household refrigerator and home freezer (335222) | 1.0500 |
| Confectionery from purchased chocolate (311330) | 1.0515 |
| Tortilla (311830) | 1.0537 |
| Breakfast cereal (311230) | 1.0562 |
| Iron and steel mills and ferroalloy (331110) | 3.3581 |
| Nonferrous metal rolling, drawing extruding and alloying (331490) | 3.3958 |
| Primary smelting and refining of nonferrous metal (331419) | 3.4186 |
| Carbon black (325182) | 3.4193 |
| Steel product from purchased steel (331200) | 3.4500 |
| Clay and nonclay refractory (32712B) | 3.4648 |
| Ground or treated mineral and earth (327992) | 3.4857 |
| Printing ink (325910) | 3.4880 |
| Synthetic dye and pigment (325130) | 3.5183 |
| Pulp mills (322110) | 3.5506 |
| Plastics material and resin (325211) | 3.5712 |
| Copper rolling, drawing, extruding and alloying (331420) | 3.6109 |
| Alkalies and chlorine (325181) | 3.6112 |
| Carbon and graphite product (335991) | 3.7484 |
| Fertilizer (325310) | 3.7617 |
| Alumina refining and primary aluminum (33131A) | 3.8144 |
| Other basic organic chemical (325190) | 3.8529 |
| Secondary smelting and alloying of aluminum (331314) | 4.0637 |
| Primary smelting and refining of copper (331411) | 4.3547 |
| Petrochemical (325110) | 4.6511 |

Notes: Tabulated for manufacturing only. Six-digit US I-O industry codes are in parentheses.

Table 2: Correlation of Upstreamness with other Industry Characteristics

| Industry variable | Correlation with upstreamness |
|---|---|
| *Capital-intensity* | |
| Log (Capital Stock / Total Workers) | 0.433*** |
| Log (Capital Expenditures / Payroll) | 0.395*** |
| Log (Capital Stock / Payroll) | 0.446*** |
| *Skill-intensity* | |
| Log (Non-Production Workers / Total Workers) | −0.046 |
| Non-production payroll / Payroll | −0.184*** |

Notes: *** indicates significance at the 1% level. Pearson linear correlations are reported. All industry variables in the left-most column are 1996-2005 averages, calculated from the NBER-CES Database for 279 manufacturing industries.

among the most downstream of industries, with almost all of their output going directly to the end-user. On the other hand, the most upstream industries tend to be involved in the processing of raw materials.[8]

A natural question to ask is how upstreamness correlates with other industry characteristics, particularly factor intensities. We construct measures of the latter using the most recent version of the NBER-CES Manufacturing Database (Bartelsman et al. 2009). Table 2 reports the correlations between our industry upstreamness variable and several alternative measures that have been commonly used to capture skill- and capital-intensity. Specifically, we take the average annual value between 1996-2005 for each of these industry variables as calculated from the NBER-CES Database.[9] Note however that given the nature of the NBER-CES Database, we are able to construct these variables only for the manufacturing sector, so the correlations are based only on the subset of 279 manufacturing industries.

From Table 2, we find that upstreamness is associated with greater physical capital-intensity, with this correlation being highly significantly regardless of how capital-intensity is calculated, namely log(Capital stock per worker), log(Capital expenditures over payroll),

[8] For the 426 industries, the correlation between upstreamness calculated with the open-economy and inventories corrections and upstreamness calculated without these corrections is a relatively high 0.89.

[9] Although the NBER-CES Database provides information for NAICS industries, these are very similar to the 2002 I-O codes. We applied the straightforward mapping between the two classification systems available on the BEA website. Note also that the NAICS system came into use in 1997, so this raises some concerns over concordance accuracy for pre-1997 data. It turns out however that the correlation between the 1996-2005 and 1997-2005 average values for each of these industry variables is very high (in excess of 0.99), so that it makes little difference if we drop the 1996 data.

or log(Capital stock over payroll). On the other hand, upstreamness is negatively correlated with skill-intensity, but more weakly so: The correlation is only significant when using the non-production worker share of total payroll to measure skill-intensity, and not significant when using the log of the non-production worker share of total employment.[10]

# 4 Upstreamness in Other Countries

The upstreamness measure is most likely to be useful if its ranking is stable across countries. In practice, stability is somewhat difficult to verify because national I-O tables differ in their product/industry classifications and the level of aggregation employed. Fortunately, there have been some efforts to collect and produce I-O tables that are consistent across countries. The OECD STAN database contains easily accessible I-O tables for many countries in a reasonably well-concorded fashion. A subset of the STAN tables were submitted by Eurostat, the statistics office of the European Union. We employ the STAN data for a subset of 16 EU countries that share an exact aggregation of the data for 2005.[11] These Eurostat tables contain 41 sectors, 13 of which are in manufacturing. As the rest of our paper relies on US data, we also check whether upstreamness calculated from the US table in the STAN database is highly correlated with the European measures. Bear in mind however that different national industry definitions mean that the US data is aggregated differently in the STAN database than in the European data we employ. In particular, three industries that are reported for the European countries are not reported for the US.

We calculate the upstreamness measure for each individual country, following the methodology described in Section 3. To verify the consistency of industry upstreamness across countries, we conduct a Spearman rank correlation test among all country pairs in the sample. These results are reported in Table 3. The rank correlation is always large and positive; in all country pairs, this is significantly different from zero at a p-value of 0.01. In particular, the US measures yield industry rankings that are consistent with those from the European data. A useful point to note is that the correlations tend to be slightly lower for small countries where trade features as a large percentage of output, since in such countries, the open-economy adjustment would matter more. Luxembourg is a clear outlier in this regard, in that the correction for trade generates an upward shift in its measures of upstreamness relative to what is observed in less trade-dependent countries.[12]

[10] See Fally (2011) for correlations with other industry variables widely used in empirical work.

[11] The included countries are: Austria, Belgium, the Czech Republic, Denmark, Estonia, Finland, Germany, Greece, Hungary, Italy, Luxembourg, the Netherlands, Portugal, Slovakia, Slovenia and Spain.

[12] For example, Luxembourg's Finance & Insurance sector has an upstreamness measure of 22.28. Only Luxembourg has outliers so large that they affect measures of central tendency across the European sample.

Table 3: Rank Correlations of Industry Upstreamness across Countries

| | USA | AUT | BEL | CZE | DEU | DNK | ESP | EST | FIN | GRC | HUN | ITA | LUX | NLD | PRT | SVK | SVN |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| USA | 1.00 | | | | | | | | | | | | | | | | |
| AUT | 0.78 | 1.00 | | | | | | | | | | | | | | | |
| BEL | 0.70 | 0.84 | 1.00 | | | | | | | | | | | | | | |
| CZE | 0.60 | 0.72 | 0.77 | 1.00 | | | | | | | | | | | | | |
| DEU | 0.78 | 0.85 | 0.86 | 0.80 | 1.00 | | | | | | | | | | | | |
| DNK | 0.75 | 0.75 | 0.86 | 0.72 | 0.83 | 1.00 | | | | | | | | | | | |
| ESP | 0.79 | 0.87 | 0.80 | 0.80 | 0.86 | 0.78 | 1.00 | | | | | | | | | | |
| EST | 0.65 | 0.65 | 0.71 | 0.71 | 0.82 | 0.82 | 0.71 | 1.00 | | | | | | | | | |
| FIN | 0.80 | 0.84 | 0.77 | 0.78 | 0.80 | 0.80 | 0.86 | 0.65 | 1.00 | | | | | | | | |
| GRC | 0.75 | 0.83 | 0.85 | 0.72 | 0.91 | 0.82 | 0.84 | 0.81 | 0.75 | 1.00 | | | | | | | |
| HUN | 0.69 | 0.76 | 0.70 | 0.90 | 0.84 | 0.69 | 0.81 | 0.69 | 0.82 | 0.72 | 1.00 | | | | | | |
| ITA | 0.81 | 0.81 | 0.80 | 0.79 | 0.87 | 0.74 | 0.86 | 0.72 | 0.83 | 0.85 | 0.79 | 1.00 | | | | | |
| LUX | 0.66 | 0.71 | 0.78 | 0.56 | 0.75 | 0.72 | 0.61 | 0.70 | 0.60 | 0.84 | 0.55 | 0.74 | 1.00 | | | | |
| NLD | 0.75 | 0.88 | 0.85 | 0.76 | 0.86 | 0.86 | 0.82 | 0.70 | 0.90 | 0.84 | 0.79 | 0.86 | 0.71 | 1.00 | | | |
| PRT | 0.74 | 0.90 | 0.85 | 0.85 | 0.87 | 0.77 | 0.94 | 0.72 | 0.82 | 0.88 | 0.79 | 0.88 | 0.66 | 0.86 | 1.00 | | |
| SVK | 0.56 | 0.77 | 0.79 | 0.75 | 0.82 | 0.68 | 0.78 | 0.73 | 0.63 | 0.78 | 0.72 | 0.69 | 0.57 | 0.68 | 0.79 | 1.00 | |
| SVN | 0.62 | 0.83 | 0.84 | 0.90 | 0.84 | 0.73 | 0.86 | 0.76 | 0.76 | 0.78 | 0.79 | 0.81 | 0.61 | 0.78 | 0.92 | 0.82 | 1.00 |

Notes: All Spearman rank correlations are significantly different from zero at the 1% level.

Table 4: Spearman rank correlation with US upstreamness

| AUS | BRA | CAN | CHL | CHN | FRA | GBR | IDN | IND | ISR |
|---|---|---|---|---|---|---|---|---|---|
| 0.78 | 0.79 | 0.87 | 0.65 | 0.45 | 0.75 | 0.84 | 0.63 | 0.67 | 0.64 |
| JPN | KOR | MEX | NOR | NZL | POL | ROU | SWE | TUR | |
| 0.56 | 0.72 | 0.62 | 0.80 | 0.81 | 0.73 | 0.66 | 0.74 | 0.72 | |

Notes: Sample includes countries in the OECD-STAN database with industry classifications that are imperfectly concorded with those in Table 3. All Spearman rank correlations are significantly different from zero at the 1% level, except USA-CHN, which is significant at the 5% level.

We also check the joint correlation of upstreamness across all 16 European countries through a principal component analysis, and find that 76 percent of the total variation in the measure is captured by a single component. Thus, not only are the measures correlated among pairs of countries, the measures are *jointly* correlated to a very high degree. Moreover, the correlation of US upstreamness with the principal component of the European measures is 0.82.

The variation of our upstreamness measure in the European data is also largely consistent with the range of values reported earlier in Table 1. In the European countries other than Luxembourg, we find a mean upstreamness of 2.50, and a standard deviation of 0.84. The mean upstreamness for industries across European countries ranges from 1.09 (Health and social work) to 3.92 (Iron and steel). In sum, the European evidence gives us great confidence that the industry measures are stable across countries, at least at the higher level of aggregation reported in the STAN database.

As noted, concordance issues make it difficult to summarize all the relationships among country measures of upstreamness in the OECD data. Nonetheless, we also conducted the Spearman rank correlation test against the wider set countries with I-O tables in the OECD STAN database, even though this typically reduces the number of industries which we can confidently match across countries. The countries with available data from the period 2002-2005 include all OECD member states (excluding Switzerland and Iceland), as well as several large non-member countries (Brazil, China, India, Indonesia and Romania). The pairwise correlations among all these country measures (including those from the previous set of countries in Table 3) are positive and statistically significant in the vast majority of cases. In order to summarize the information succinctly, we report correlations between the US measure of upstreamness calculated from the OECD STAN data and that for these additional countries in Table 4. These correlations are all reassuringly high in spite of the above-mentioned concordance issues.[13]

# 5 Application to Trade

We briefly explore how our measure of industry upstreamness, specifically that based on the more disaggregate 2002 US I-O Tables, can provide some new perspectives on trade patterns at the country level. In particular, with this new measure, we are now equipped to describe a country's average position in global production chains, namely whether the country tends on average to be an exporter in relatively upstream versus downstream industries.

Toward this end, we calculate a summary measure of the upstreamness of a country's exports as follows. Data on world trade flows at the Harmonized System six-digit (HS6) level are taken from the BACI dataset.[14] BACI draws originally on the UN Comtrade database, but applies a procedure to harmonize and clean the data to reconcile trade flows reported by exporting and importing countries. We map the trade flows from HS6 to US I-O 2002 categories using a concordance provided by the BEA. We then take a weighted average of industry upstreamness values for each country, using the total exports by the country in the respective industries as weights. Naturally, this assumes that the US measures of upstreamness provide a good description of production line position in other countries as well, but as we have seen in Section 3, this appears to be a reasonable starting point.

In what follows, we consider trade flows from 2002 for a core sample of 181 countries.[15] Constructing country upstreamness as described above, we obtain a mean value of export upstreamness of 2.30 with a standard deviation of 0.58. If attention is restricted to manufacturing trade flows, this mean country upstreamness falls to 2.05, with a standard deviation of 0.49. This drop reflects the fact that many primary and resource-extracting industries tend to enter production processes at relatively upstream stages.[16]

Looking beyond these broad averages, Table 5 reports the mean values of export upstreamness by country income groups. We split the countries in our sample into quartiles, as determined by the mean log real GDP per capita between 1996-2005, calculated from the Penn World Tables, Version 7.0 (Heston et al. 2011). At first glance, taking into consider-

[13] All of the upstreamness measures we calculate for country-industry pairs in the OECD database are available from the authors.

[14] Available at http://www.cepii.fr/anglaisgraph/bdd/baci.htm

[15] This consists of the 181 countries for which the export upstreamness measure could be constructed, and for which data on real GDP per capita for 1996-2005 was available in the Penn World Tables, Version 7.0. We merged Belgium and Luxembourg as the BACI do not report separate trade flows for the two countries.

[16] The mean value of our upstreamness measure for the 30 industries related to agriculture, forestry and mining (I-O codes starting with '1' or '21') is 2.84, compared to the mean upstreamness of 2.10 for the 279 manufacturing industries (I-O codes starting with '3').

Table 5: Upstreamness of Exports by Country Income Quartiles

| | All | | Manufacturing | |
|---|---|---|---|---|
| Income quartile | Mean | S.D. | Mean | S.D. |
| Bottom | 2.41 | 0.69 | 2.03 | 0.60 |
| 2nd | 2.30 | 0.60 | 1.98 | 0.48 |
| 3rd | 2.23 | 0.55 | 2.11 | 0.51 |
| Top | 2.26 | 0.45 | 2.10 | 0.34 |

Notes: Countries are grouped into income quartiles based on the average log PPP-adjusted GDP per capita over 1996-2005, from the Penn World Tables, Version 7.0. The average upstreamness of country exports and its standard deviation within each quartile is reported under the first set of columns labeled "All". The second set of columns restricts the calculation to manufacturing exports only.

ation all trade flows, the export activities of poorer countries appear to be in slightly more upstream industries than that of richer countries. However, when we focus in on manufacturing trade flows alone, no simple relationship between country per capita GDP and export upstreamness is evident. This is not entirely surprising given that we have seen that diverse manufacturing industries can feature similar values of upstreamness. Recall for instance that automobiles and footwear both rank among the five most downstream industries.

More interestingly, the standard deviation of export upstreamness within each country quartile decreases as the mean income level rises. Countries in the top quartile are thus more similar in terms of the average position they occupy in global production lines, while there is much more variation across poorer countries on this dimension.[17] To give an example, consider Bangladesh and Tajikistan, two countries with a similarly low level of per capita income. Although both countries are in the bottom income quartile of our sample, they are at opposite ends of the spectrum in terms of export upstreamness. Bangladesh ranks among the five most downstream countries in terms of its manufacturing exports (country upstreamness = 1.26), due to its position as a major exporter of apparel, a good that tends to be sold directly to end-consumers. Tajikistan instead ranks among the five most upstream countries (country upstreamness = 3.53), as processed alumina takes up the lion's share of its exports. Once again, there does not appear to be a simple uniform story that connects a country's income level to its average production line position.

Building on this discussion, we examine the correlations between export upstreamness and various country characteristics more systematically in Table 6. We stress that our objective here is not to establish causality or investigate particular mechanisms, but simply

[17] A similar conclusion is reached if we consider the coefficient of variation instead.

to uncover interesting patterns that relate to a country's average production line position. Panel A in Table 6 reports regression findings in which country upstreamness based on all exports is the dependent variable, while Panel B reports the corresponding findings when upstreamness is calculated for manufacturing exports only. We use explanatory variables that are from standard sources of cross-country data; where possible, we have calculated these as averages over 1996-2005.

In Column 1, we verify that the simple bivariate correlation between country upstreamness and log real GDP per capita (from the Penn World Tables) is not statistically significant. We find much more interesting results in Columns 2-4 where we introduce variables related to country institutions, namely: (i) a rule of law index from Kaufmann et al. (2011), that is often used as an indicator of the strength of contracting institutions; and (ii) the ratio of private credit to GDP from Beck et al. (2010), reflecting the level of financial development in the economy. The negative partial correlations obtained here imply that better rule of law and stronger financial development are associated at the country level with a basket of exports that is relatively more downstream in terms of production line position.

Column 5 explores whether factor endowments have a role to play in determining a country's export upstreamness. We include a measure of log physical capital per worker, calculated from the Penn World Tables using the perpetual inventory method in Hall and Jones (1999), as well as the average years of schooling in the population aged 15 and over from Barro and Lee (2010). The findings here suggest that the negative correlation between country upstreamness and financial development is a particularly robust one; that for country rule of law in contrast becomes imprecisely estimated in both panels. Moreover, there appears to be some potential role for factor endowments in explaining a country's average production line position, as human capital is associated with more downstream exports. This last finding nevertheless needs to be taken with a pinch of salt, as this correlation with years of schooling is no longer significant in the lower panel that focuses on manufacturing trade flows.[18]

We have also examined the average upstreamness of imports by country using the same approach as that taken for exports. First, we find that the calculated country import upstreamness variable has a lower mean (2.10 if we consider all commodities and 2.01 for manufacturing only) and a smaller standard deviation (0.25 for all commodities and 0.22 for manufacturing only), as compared to export upstreamness. The Pearson correlation between import and export upstreamness across countries is small and not significant if we consider all commodities (0.07), but becomes larger and significant at the 1% level if we restrict our sample to manufacturing industries (0.20). In terms of cross-country regressions (as in Ta-

[18] All our results are similar when controlling for a measure of openness, namely exports plus imports over GDP, from the Penn World Tables. This variable itself has little explanatory power for export upstreamness.

Table 6: Export Upstreamness and Country Characteristics

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| *Panel A: Country Upstreamness, All Exports (2002)* | | | | | |
| Log (Real GDP per capita) | −0.035 | 0.146*** | 0.100** | 0.156** | 0.083 |
| | (0.032) | (0.054) | (0.047) | (0.060) | (0.142) |
| Rule of Law | | −0.313*** | | −0.164* | −0.029 |
| | | (0.070) | | (0.091) | (0.103) |
| Private Credit / GDP | | | −0.585*** | −0.404*** | −0.437*** |
| | | | (0.123) | (0.128) | (0.136) |
| Log (Capital per worker) | | | | | 0.156 |
| | | | | | (0.131) |
| Years of Schooling | | | | | −0.085*** |
| | | | | | (0.031) |
| $N$ | 181 | 181 | 151 | 151 | 120 |
| $R^2$ | 0.01 | 0.11 | 0.09 | 0.11 | 0.15 |
| *Panel B: Country Upstreamness, Manufacturing Exports (2002)* | | | | | |
| Log (Real GDP per capita) | 0.031 | 0.112** | 0.115*** | 0.124** | 0.056 |
| | (0.028) | (0.053) | (0.042) | (0.061) | (0.140) |
| Rule of Law | | −0.140** | | −0.027 | 0.045 |
| | | (0.068) | | (0.088) | (0.094) |
| Private Credit / GDP | | | −0.312*** | −0.282** | −0.274** |
| | | | (0.105) | (0.111) | (0.116) |
| Log (Capital per worker) | | | | | 0.053 |
| | | | | | (0.118) |
| Years of Schooling | | | | | −0.026 |
| | | | | | (0.027) |
| $N$ | 181 | 181 | 151 | 151 | 120 |
| $R^2$ | 0.01 | 0.04 | 0.06 | 0.06 | 0.05 |

Notes: Robust standard errors in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% levels respectively. All right-hand side variables are averages over annual data from available years between 1996-2005.

ble 6), we find no significant effect of any of the country variables used above except for GDP per capita which is associated with a negative coefficient (results available on request). The relevance of these country variables in explaining trade patterns is thus specific to the supply-side, and does not appear to be driven by differences in the composition of demand.

So far we have focused on studying the effects of country characteristics on the average upstreamness of their exports. An alternatively, widely-used approach in the literature consists in exploiting the full cross-country and cross-industry variation in trade flow data by running specifications of the form

$$\log X_{ci} = \alpha_c + \beta_i + \sum_l \gamma_l \times V_{c,l} \times U_i + \sum_{l,m} \phi_{l,m} \times V_{c,l} \times S_{i,m} + \varepsilon_{ci}.$$

The dependent variable is now the log of exports from country $c$ in industry $i$ in the year 2002 (again from the BACI data source), at the US I-O 2002 industry level. To examine whether the production line position of an industry influences the cross-country, cross-industry pattern of trade, we include on the right-hand side industry upstreamness, $U_i$, interacted with a set of country characteristics, $V_{c,l}$, which are indexed by $l$. The $\alpha_c$'s and $\beta_i$'s are full sets of country and industry fixed effects respectively, while $\varepsilon_{ci}$ is a standard noise term; in practice, we will report robust standard errors. In some later specifications, we will further control for the interaction between country variables and other industry characteristics, $S_{i,m}$, indexed by $m$, to capture other potential determinants of the pattern of trade that arise from a country's ability to facilitate specialization in industries with particular production requirements.

Column 1 of Table 7 reports a baseline specification in which we interact all the country variables considered previously in Table 6 with industry upstreamness, to explore whether these help to explain trade patterns. The interaction between country rule of law and industry upstreamness is highly significant at the 1% level, indicating that countries with better rule of law tend to export more in industries that are relatively downstream. It does appear too that there is a similar tendency for countries with better financial development to export more in downstream industries, but this coefficient is not statistically significant. Exploiting the fact that we now have more data points, Column 2 includes an interaction between the square of log per capita income and $U_i$. We do uncover evidence of a non-linear effect of country income on patterns of trade, with the coefficients suggesting that countries at very high or very low levels of per capita income would tend to feature more upstream exports.[19] Of note, the effects of country rule of law and financial development remain similar in this specification. Column 2 further reveals that countries that are relatively abundant in

[19] We do not obtain significant results when including the square of log real per capita GDP in the earlier Table 6 regressions.

physical capital tend to be more active in exporting in upstream industries.

In Columns 3 and 4, we re-run these first two specifications focusing only on trade flows for the manufacturing sector. The results here are very similar, particularly for the effects of a country's rule of law and physical capital endowment, in influencing the relative production line position of its exports. However, the coefficient of the interaction involving financial development is now essentially zero and very imprecisely estimated. In Column 5, we include further interaction terms between country factor endowments and industry factor intensities (for both human and physical capital), to control for endowment-based motives for trade of a Heckscher-Ohlin flavor following Romalis (2004). This does not affect our earlier findings regarding how country variables interact with industry upstreamness. (We have used the log capital stock per worker and log non-production share of total employment variables in these regressions, but the results are very similar when using the other factor intensity variables in Table 2.) Lastly, we cluster our standard errors by exporting country in Column 6, but the correlations we have found continue to remain significant at least at the 10% level.[20]

Summarizing our findings from Tables 6 and 7, country institutions such as the rule of law and financial development appear to be correlated with the relative production line position of a country's exports. There is some evidence too that factor endowments, namely physical capital and skill abundance, do matter in this regard. We view these findings as motivation for possible future theoretical work to understand how underlying country conditions interact with production line position in determining patterns of specialization and trade.

# 6 Conclusion

We have developed and constructed a measure of industry upstreamness in this short note. The empirical applications which we presented in Section 5, though preliminary in their nature, suggest that this is an industry attribute that warrants further attention particularly in this age of cross-border production fragmentation. We have started exploring these potential research directions ourselves in our separate work. Fally (2011) for example has explored issues related to production line position, its evolution over time, and its implications for comparative advantage, using more detailed time-series, cross-country, cross-industry variation in trade flows. Taking contracting issues seriously, Antràs and Chor (2011) seek to understand how cross-border firms would seek to organize themselves along production chains, vis-à-vis the integration versus outsourcing decision.

[20] All our results in Table 7 are qualitatively similar if we omit trade flows related to the extraction and processing of petroleum (results available on request).

Table 7: Industry Upstreamness and the Pattern of Trade

| Dep. variable: Log trade flows | (1) All | (2) All | (3) Manuf. | (4) Manuf. | (5) Manuf. | (6) Manuf. |
|---|---|---|---|---|---|---|
| Log (Real GDP per capita) × $U_i$ | −0.036 | −0.675*** | −0.069 | −0.542*** | −0.543*** | −0.543* |
| | (0.050) | (0.209) | (0.049) | (0.205) | (0.204) | (0.277) |
| (Log (Real GDP per capita))$^2$ × $U_i$ | | 0.036*** | | 0.026** | 0.026** | 0.026 |
| | | (0.011) | | (0.011) | (0.011) | (0.016) |
| Rule of Law × $U_i$ | −0.105*** | −0.148*** | −0.079*** | −0.110*** | −0.110*** | −0.110* |
| | (0.030) | (0.032) | (0.029) | (0.032) | (0.032) | (0.061) |
| (Private Credit / GDP) × $U_i$ | −0.059 | −0.072 | 0.010 | 0.001 | 0.001 | 0.001 |
| | (0.048) | (0.048) | (0.047) | (0.047) | (0.047) | (0.075) |
| Log (Capital per worker) × $U_i$ | 0.048 | 0.093** | 0.142*** | 0.176*** | 0.194*** | 0.194*** |
| | (0.044) | (0.047) | (0.043) | (0.045) | (0.045) | (0.074) |
| Years of Schooling × $U_i$ | −0.009 | −0.001 | −0.007 | 0.000 | 0.000 | 0.000 |
| | (0.009) | (0.009) | (0.009) | (0.009) | (0.009) | (0.016) |
| Log (Capital per worker) × $\log(k/l)_i$ | | | | | −0.040*** | −0.040* |
| | | | | | (0.012) | (0.024) |
| Years of Schooling × $\log(s/l)_i$ | | | | | 0.073*** | 0.073*** |
| | | | | | (0.010) | (0.027) |
| Industry fixed effects? | Yes | Yes | Yes | Yes | Yes | Yes |
| Country fixed effects? | Yes | Yes | Yes | Yes | Yes | Yes |
| $N$ | 32822 | 32822 | 29215 | 29215 | 29215 | 29215 |
| $R^2$ | 0.77 | 0.77 | 0.80 | 0.80 | 0.80 | 0.80 |

Notes: ***, **, and * denote significance at the 1%, 5%, and 10% levels respectively. Robust standard errors are reported, except in Column (6) where these are clustered by exporting country instead. The dependent variable in Columns (1)-(2) is log trade flows, while that in the remaining columns restricts to log manufacturing trade flows. In Columns (5)-(6), the industry capital intensity measure, $\log(k/l)_i$, is the log capital stock per worker, while the skill intensity measure, $\log(s/l)_i$, is the log non-production worker share of total employment. All right-hand side variables are constructed as averages over annual data from available years between 1996-2005.

# References


Antràs, Pol, and Davin Chor. 2011. "Organizing the Global Value Chain." Mimeo.

Barro, Robert, and Jong Wha Lee. 2010. "A New Data Set of Educational Attainment in the World, 1950-2010." NBER Working Paper No. 15902.

Bartelsman, Eric J., Randy A. Becker, and Wayne B. Gray. 2009. NBER-CES Manufacturing Industry Database.

Beck, Thorsten, Asli Demirgüç-Kunt, and Ross Levine. 2010. "A New Database on Financial Development and Structure." World Bank dataset.

Dietzenbacher, Erik. 1997. "In Vindication of the Ghosh Model: A Reinterpretation as a Price Model." *Journal of Regional Science*, 37: 629-651.

Fally, Thibault. 2011. "On the Fragmentation of Production in the U.S." Mimeo.

Ghosh, Ambica. 1958. "Input-Output Approach in an Allocation System." *Economica*, 25: 58-64.

Hall, Robert E., and Charles I. Jones. 1999. "Why Do Some Countries Produce So Much More Output Per Worker Than Others?" *Quarterly Journal of Economics*, 114: 83-116.

Heston, Alan, Robert Summers, and Bettina Aten. 2011. "Penn World Table Version 7.0." Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania.

Horowitz, Karen J., and Mark A. Planting. 2009. "Concepts and Methods of the Input-Output Account." Bureau of Economic Analysis documentation.

Kaufmann, Daniel, Aart Kraay, and Massimo Mastruzzi. 2011. "World Governance Indicators, 1996-2010." World Bank dataset.

Krugman, Paul. 1995. "Growing World Trade: Causes and Consequences." *Brookings Papers on Economic Activity*, 1: 327-377.

Miller, Ronald E., and Peter D. Blair. 2009. *Input-Output Analysis: Foundations and Extensions.* Second Edition. Cambridge University Press.

Romalis, John. 2004. "Factor Proportions and the Structure of Commodity Trade." *American Economic Review*, 94: 67-97.

# A Appendix

This Appendix provides the proof of our key result on the equivalence of the two separately-defined measures of industry upstreamness, as well as the proofs of the two additional interpretations. It further derives the open-economy adjustment, as well as that for the treatment of inventories. Finally, it provides details on the two-country, two-industry example developed in the text. The notation follows that in our main text.

## A.1 Proof of Proposition 1

Recall first that $U_{2i}$ is defined recursively by

$$U_{2i} = 1 + \sum_{j=1}^{N} \frac{d_{ij} Y_j}{Y_i} U_{2j}. \tag{10}$$

Multiply both sides of (10) by $Y_i$ to obtain

$$U_{2i} Y_i = Y_i + \sum_{j=1}^{N} d_{ij} U_{2j} Y_j.$$

Defining $P_i = U_{2i} Y_i$ for all $i \in \{1, 2, ..., N\}$, we have

$$P_i = Y_i + \sum_{j=1}^{N} d_{ij} P_j. \tag{11}$$

Let $P$ be the column vector whose $i$-th entry is $P_i$. Also, let $D$ denote the square matrix whose $(i, j)$-th entry is $d_{ij}$, which is the amount of input $i$ needed to produce one dollar of $j$. Stacking up the $P_i$'s in (11) into column-vector form, we have: $P = Y + DP$, where $Y$ is the column vector with $Y_i$ as its $i$-th entry.

Solving for $P$ leads to

$$P = [I - D]^{-1} Y.$$

But as discussed in Section 2, the $i$-th row of $[I - D]^{-1} Y$ is precisely equal to the numerator of $U_{1i}$. Since $P_i / Y_i = U_{2i}$, it follows that $U_{2i} = U_{1i}$.

## A.2 Proof of Two Economic Interpretations

To show that

$$U_i = \frac{1}{Y_i} \sum_{j=1}^{N} \frac{\partial Y_i}{\partial d_{jj}},$$

start from the basic identity that decomposes output for industry $i$ into its final-use value and that which goes to intermediate uses:

$$Y_i = F_i + \sum_{j=1}^{N} d_{ij} Y_j. \tag{12}$$

Differentiating (12) yields

$$\begin{aligned} \frac{\partial Y_i}{\partial d_{ii}} &= Y_i + \sum_{k=1}^{N} d_{ik} \frac{\partial Y_k}{\partial d_{ii}}, \text{ and} \\ \frac{\partial Y_i}{\partial d_{jj}} &= \sum_{k=1}^{N} d_{ik} \frac{\partial Y_k}{\partial d_{jj}} \text{ for } j \neq i. \end{aligned}$$

Summing over these partial derivatives and dividing by $Y_i$, note that we can write

$$\frac{1}{Y_i} \sum_{j=1}^{N} \frac{\partial Y_i}{\partial d_{jj}} = 1 + \sum_{k=1}^{N} \left( \frac{d_{ik} Y_k}{Y_i} \right) \frac{1}{Y_k} \sum_{j=1}^{N} \frac{\partial Y_k}{\partial d_{jj}}.$$

But note that this is the same recursive equation that defines $U_{2i}$ in (10), so this establishes that $U_i = \frac{1}{Y_i} \sum_{j=1}^{N} \frac{\partial Y_i}{\partial d_{jj}}$.

Next, to show that, holding constant the allocation matrix $\Delta$, we have

$$U_i = \sum_{j=1}^{N} \frac{\partial Y_j}{\partial V_i},$$

we begin from the identity equating industry revenue $Y_i$ to its outlays in terms of primary factors $V_i$ and input purchases:

$$Y_i = V_i + \sum_{j=1}^{N} d_{ji} Y_i = V_i + \sum_{j=1}^{N} \frac{d_{ji} Y_i}{Y_j} Y_j,$$

Denoting by $\delta_{ji}$ the $(j,i)$-th element of the allocation matrix $\Delta$, namely $d_{ji} Y_i / Y_j$, note that we can iterate this equation in a manner analogous to the derivation of (2) in the main text:

$$Y_i = V_i + \sum_{j=1}^{N} \delta_{ji} V_j + \sum_{j=1}^{N} \sum_{k=1}^{N} \delta_{jk} \delta_{ki} V_j + \sum_{j=1}^{N} \sum_{k=1}^{N} \sum_{l=1}^{N} \delta_{jk} \delta_{kl} \delta_{li} V_j + ...$$

Stacking these equations for all industries $i \in \{1, 2, ..., N\}$, we can write the system compactly in matrix notation:

$$Y = V + \Delta' V + \left(\Delta'\right)^2 V + \left(\Delta'\right)^3 V + ....$$

Provided that $\sum_{j=1}^{N} \delta_{ij} < 1$ for all $i$ (a natural assumption), this expression converges to

$$Y = \left[I - \Delta'\right]^{-1} V = \left(\left[I - \Delta\right]^{-1}\right)' V.$$

Differentiating this system with respect to $V_i$ and adding over all $j$ then delivers

$$\sum_{j=1}^{N} \frac{\partial Y_j}{\partial V_i} = i\text{-th element of } [I - \Delta]^{-1} \mathbf{1}.$$

where $\mathbf{1}$ is a column vector of $N$ one's. But remember from Section 2 that this is identical to $U_2$ and thus

$$\sum_{j=1}^{N} \frac{\partial Y_j}{\partial V_i} = U_{2i} = U_i.$$

Note also that this derivation emphasizes that our measure of upstreamness is equal to the row sum of the so-called Ghosh inverse matrix $[I - \Delta]^{-1}$, which in turn is a standard measure of total forward linkages in Input-Output models (see Miller and Blair, 2009, p. 558).

## A.3 Open-Economy Adjustment

Recall that we calculate our upstreamness measure for all $N$ industries via the formula $[I - \Delta]^{-1}\mathbf{1}$, with $\mathbf{1}$ being a column vector of $N$ one's, and $\Delta$ being the square matrix whose $(i, j)$-th entry is $d_{ij}Y_j/Y_i$ in the closed-economy setting. In the open-economy setting, we therefore need to determine the correction that needs to be applied in order for the entries of the matrix $\Delta$ to continue to reflect the share of domestic output from $i$ that is purchased by industry $j$ as inputs, regardless of whether the purchasing industry is located at home or abroad.

With trade, the basic output identity for each industry $i$ now becomes

$$\begin{aligned} Y_i &= F_i + Z_i + X_i - M_i \\ &= F_i + X_{Fi} - M_{Fi} + Z_i + X_{Zi} - M_{Zi}, \end{aligned}$$

where $F_i$ and $Z_i$ denote the value of output produced domestically in industry $i$ that goes respectively towards final uses and intermediate input uses *in the home economy.* $X_i$ and $M_i$ denote total exports and imports of industry $i$. These in turn can be broken down into exports and imports that go to final uses ($X_{Fi}$ and $M_{Fi}$ respectively), and exports and imports that are used as inputs in the production of other goods ($X_{Zi}$ and $M_{Zi}$).

Note that

$$Z_i + X_{Zi} - M_{Zi} = \sum_{j=1}^{N} \left( d_{ij}Y_j + X_{ij} - M_{ij} \right),$$

where the sum is taken over industries $j$ that purchase inputs of $i$. $X_{ij}$ and $M_{ij}$ refer respectively to the exports and imports from industry $i$ that are purchased for intermediate input use specifically in industry $j$. In the open-economy, let $\delta_{ij}$ denote the share of $i$'s domestic production that is purchased directly by industry $j$ (both at home or abroad). $\delta_{ij}$ is thus given by

$$\delta_{ij} = \frac{d_{ij}Y_j + X_{ij} - M_{ij}}{Y_i}, \tag{13}$$

This takes into account the fact that in the US I-O Tables, the final-use and intermediate-use values reported do not distinguish between goods/intermediates that are produced domestically versus that which is imported.[21]

One problem with taking (13) directly to the data is that $X_{ij}$ and $M_{ij}$ are typically not observed. To make progress, we argue that it is reasonable to assume that the share of industry $i$'s output used in industry $j$ (at home and abroad) be identical to the share of industry $i$'s exports (imports) that are used by industry $j$ producers, namely

$$X_{ij} = \delta_{ij} X_i$$

and

$$M_{ij} = \delta_{ij} M_i.$$

This is precisely Assumption 1 in our main paper. Substituting these expressions into (13), straightforward manipulation leads to

$$\delta_{ij} = \frac{d_{ij} Y_j}{Y_i - X_i + M_i}.$$

We therefore implement the open-economy adjustment by replacing $d_{ij} Y_j / Y_i$ with the above expression for $\delta_{ij}$ for the entries of the matrix $\Delta$. This is equivalent to replacing $d_{ij}$ with

$$\hat{d}_{ij} = d_{ij} \frac{Y_i}{Y_i - X_i + M_i} \tag{14}$$

as stated in the main paper.

## A.4 Treatment of Inventories

There is one remaining item classified under final uses in the input-output tables that requires careful treatment, namely net changes to inventories. We have abstracted from this when discussing the open-economy adjustment, to avoid cluttering the notation, but it can be readily seen that a similar set of considerations is involved. The input-output matrix $D$ does not distinguish between inputs that are obtained from past inventories as opposed to new production. Taking this into account, the relevant share of $i$'s domestic production that is purchased directly by industry $j$ (both at home or abroad) is given more precisely by

$$\delta_{ij}^{inv} = \frac{d_{ij} Y_j + X_{ij} - M_{ij} + N_{ij}}{Y_i},$$

where $N_{ij}$ denotes here the net value of industry $i$ output purchased by industry $j$ for the purposes of inventorization. (The superscript 'inv' indicates that this expression for $\delta_{ij}$ explicitly spells out

[21] The Bureau of Economic Analysis does provide an accompanying "Import Matrix" with the 2002 Tables that reports final-use and intermediate-use values that come from foreign sources. However, due to the limited information on the use of imports at the industry level, the "Import Matrix" is actually constructed based on a proportionality assumption that the share of import use is the same across all final uses and industries. See Horowitz and Planting (2009) for details.

the role of net inventories.) When $N_{ij}$ is positive, this means that industry $j$ is on net increasing its inventories of input $i$; a negative $N_{ij}$ in turn indicates a net draw-down of $j$'s inventories of $i$.

To take this to the data, we once again face the problem that $N_{ij}$ is not easily observed. We therefore make the same proportionality assumption as with the open-economy correction

$$N_{ij} = \delta_{ij}^{inv} N_i,$$

where $N_i$ is the aggregate net change in inventories of output from industry $i$. In words, we assume that the share of industry $i$'s output that is purchased by industry $j$ is equal to the share of net changes of inventories of $i$ that can be attributed to the net changes made by industry $j$.

With this assumption, straightforward algebra yields

$$\delta_{ij}^{inv} = \frac{d_{ij} Y_j}{Y_i - X_i + M_i - N_i}.$$

This requires that we correct for net changes of inventories, $N_i$, in the denominator of $\delta_{ij}^{inv}$. Alternatively, as stated in the main paper, the expression for $\hat{d}_{ij}$ in (14) is valid so long as the $Y_i$ in the denominator is calculated excluding the value of net changes in inventories of $i$.

## A.5 Details of Two-Country, Two-Industry Example

Towards the end of Section 2.2, we have developed a simple example to illustrate the importance of the open-economy adjustment. Here, we provide some further details on this example. Because industry 2 produces only intermediate inputs which are entirely sold to producers in sector 1, while sector 1 produces only final goods, the matrix $D$ is composed of zeros except for the $(2,1)$-entry, i.e., in the second row and first column. Remember also that Home exports part of its production of good 1 to final consumers in Foreign, while Foreign producers of good 2 sell part of their output to Home producers in sector 1. Denote by $X_1^H$ the volume of Home net exports (or Foreign net imports) in sector 1, and by $X_2^F$ the volume of Foreign net exports (or Home net imports) in sector 2. In the data, the $(2,1)$-element of the matrix $D$ will be given by

$$d_{21}^H = \frac{Y_2^H + X_2^F}{Y_1^H},$$

at Home, and by

$$d_{21}^F = \frac{Y_2^F - X_2^F}{Y_1^F}$$

in Foreign. These correspond to the dollar amounts of sector 2's output (domestic or imported) used to produce one dollar worth of industry 1's output in each country. In practice, we do not observe whether Foreign exports entail final goods or intermediate inputs, but Assumption 1 would lead us to correctly categorize all Foreign exports as intermediate inputs.

Let us now compute upstreamness in each industry at Home and in Foreign. Consider first using our formula for $U_2$ in (4) without the open-economy adjustment. Then, at Home, we would compute

$$
\begin{aligned}
U_1^H &= 1 \\
U_2^H &= 1 + d_{21}^H \frac{Y_1^H}{Y_2^H} = 1 + \frac{Y_2^H + X_2^F}{Y_2^H} > 2,
\end{aligned}
$$

while in Foreign we would compute

$$
\begin{aligned}
U_1^F &= 1 \\
U_2^F &= 1 + d_{21}^F \frac{Y_1^F}{Y_2^F} = 1 + \frac{Y_2^F - X_2^F}{Y_2^F} < 2.
\end{aligned}
$$

Instead, applying the open-economy adjustment, we would obtain the conceptually right answers:

$$
\begin{aligned}
U_1^H &= U_1^F = 1 \\
U_2^H &= 1 + d_{21}^H \frac{Y_2^H}{Y_2^H + X_2^F} \frac{Y_1^H}{Y_2^H} = 2 \\
U_2^F &= 1 + d_{21}^F \frac{Y_2^F}{Y_2^F - X_2^F} \frac{Y_1^F}{Y_2^F} = 2.
\end{aligned}
$$

In sum, the adjusted measure delivers the correct values of upstreamness in each industry and each country, while, without the adjustment, the measure of upstreamness in industry 2 is biased upwards at Home and biased downwards in Foreign. Furthermore, the size of the bias is increasing in the value of Foreign exports to Home.