Bringing Product Space into CGE Modeling: Method and Preliminary Insights

Hans Lofgren, World Bank
Israel Osorio-Rodarte, World Bank
Martin Cicowiez, CEDLAS-Universidad Nacional de La Plata

Presentation prepared for the 16th Annual Conference on Global Economic Analysis, Shanghai, June 12-14, 2013

OUTLINE

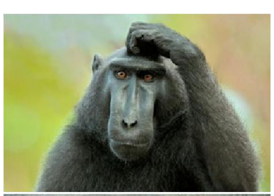
- Quick guide to Product Space (PS) Analysis
- General issues: Using PS to analyze structural change in a CGE model
- Bringing PS into MAMS
- Illustrative and potential simulations for archetype country
- Conclusions/Issues
- References
- Appendix: Technical note (PDF)

QUICK GUIDE TO PS ANALYSIS

- Trees ⇔ production sectors
- The forest ⇔ all the trees = the product space (global)
- Monkeys (sitting in or jumping between trees) ⇔ firms (producing in or moving between sectors)
- Proximity (distance) between trees ⇔ Similarity (difference) between sectors in terms of "capabilities" (human and physical capital, intermediate inputs, infrastructure needs, regulatory requirements, …)
- It is easy/hard for a monkey to jump to a nearby/distant tree ⇔ It is easy/hard for a firm to move to a similar/different sector.
- Structural transformation: process of monkeys jumping from poor to rich forest sections and climbing to tree tops ⇔ process of firms shifting from poor to rich country products and reaching rich-country quality.

All trees are not equal

A tree may be rich/poor in fruits ⇔
A sector may be dominated by rich/poor countries.



It is hard to decide: an easy jump to a nearby tree with some more bananas or a difficult jump to a distant tree with so many bananas that it is hard to count.

Fortunately, there is a scientific approach ….

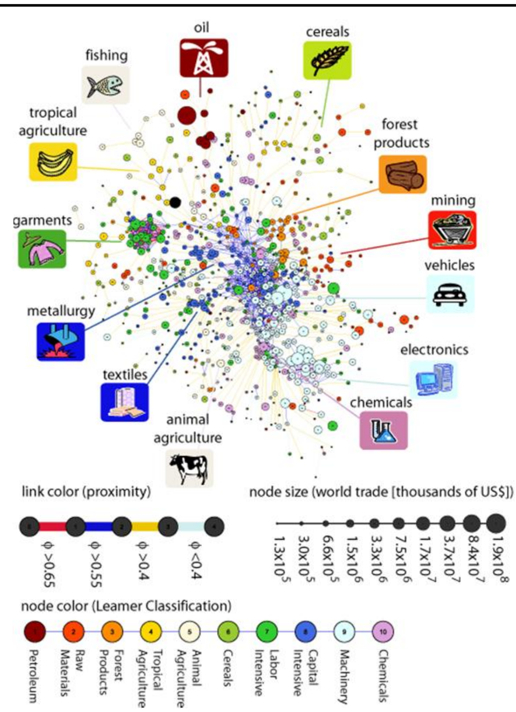
Key concepts in PS analysis

- Similarities in sectoral capabilities inferred from observing outcomes, not physical similarities of inputs or outputs.
- Data availability influences choice of concepts and their operationalizaton – trade data more abundant than production data; used as proxies for production.
- Product space and sectors: the 784 products in Comtrade's global trade records (SITC Rev. 2, 4 digits) with data for 130 countries and 30 years).
- The binary (0-1) revealed comparative advantage (RCA) à la Balassa:

$$RCA01_{c,i} = 0 \text{ or } 1$$

is used as the indicator of country performance as a productive producer of a good.

The product space as a network (Hidalgo et al. 2007)



Key concepts in PS analysis

- The *proximity* between products *i* and *i'* is the minimum of the conditional probabilities that:
  a. countries have an *RCA* (*RCA* > 1) in *i* given that they have one in *i'*; and
  b. countries have an *RCA* in *i'* given that they have one in *i*, i.e.

$$\varphi_{i,i'} = \min \left\{ \begin{matrix} P\left(RCA01_i = 1 \middle| RCA01_{i'} = 1\right), \\ P\left(RCA01_{i'} = 1 \middle| RCA01_i = 1\right) \end{matrix} \right\}$$

Table. Proximity between sectors

	agr	fish	extr	food	texclo	mnfoth	ener	chem	mme	elwat	constr	trade	hotrest	trnsp	comm	finan	busser	govser	othser
agr	1.00	0.57	0.45	0.69	0.51	0.48	0.34	0.23	0.32	0.51	0.43	0.58	0.40	0.58	0.49	0.11	0.26	0.23	0.22
fish	0.57	1.00	0.41	0.62	0.59	0.48	0.35	0.24	0.35	0.46	0.50	0.54	0.44	0.49	0.48	0.15	0.26	0.24	0.26
extr	0.45	0.41	1.00	0.38	0.34	0.28	0.26	0.23	0.33	0.40	0.30	0.35	0.20	0.39	0.35	0.26	0.23	0.26	0.18
food	0.69	0.62	0.38	1.00	0.59	0.64	0.43	0.38	0.49	0.48	0.51	0.49	0.43	0.51	0.51	0.11	0.21	0.23	0.36
texclo	0.51	0.59	0.34	0.59	1.00	0.64	0.40	0.34	0.46	0.49	0.54	0.40	0.38	0.54	0.46	0.12	0.18	0.18	0.28
mnfoth	0.48	0.48	0.28	0.64	0.64	1.00	0.52	0.48	0.65	0.47	0.50	0.37	0.33	0.43	0.50	0.07	0.15	0.22	0.46
ener	0.34	0.35	0.26	0.43	0.40	0.52	1.00	0.53	0.51	0.40	0.36	0.29	0.28	0.34	0.38	0.08	0.11	0.19	0.36
chem	0.23	0.24	0.23	0.38	0.34	0.48	0.53	1.00	0.62	0.32	0.34	0.13	0.25	0.26	0.38	0.13	0.06	0.35	0.58
mme	0.32	0.35	0.33	0.49	0.46	0.65	0.51	0.62	1.00	0.42	0.42	0.23	0.33	0.36	0.40	0.16	0.05	0.19	0.51
elwat	0.51	0.46	0.40	0.48	0.49	0.47	0.40	0.32	0.42	1.00	0.47	0.45	0.25	0.54	0.42	0.15	0.19	0.19	0.19
constr	0.43	0.50	0.30	0.51	0.54	0.50	0.36	0.34	0.42	0.47	1.00	0.29	0.40	0.49	0.50	0.12	0.16	0.28	0.44
trade	0.58	0.54	0.35	0.49	0.40	0.37	0.29	0.13	0.23	0.45	0.29	1.00	0.44	0.48	0.38	0.17	0.35	0.19	0.12
hotrest	0.40	0.44	0.20	0.43	0.38	0.33	0.28	0.25	0.33	0.25	0.40	0.44	1.00	0.28	0.42	0.13	0.33	0.35	0.38
trnsp	0.58	0.49	0.39	0.51	0.54	0.43	0.34	0.26	0.36	0.54	0.49	0.48	0.28	1.00	0.46	0.10	0.18	0.20	0.26
comm	0.49	0.48	0.35	0.51	0.46	0.50	0.38	0.38	0.40	0.42	0.50	0.38	0.42	0.46	1.00	0.17	0.23	0.35	0.38
finan	0.11	0.15	0.26	0.11	0.12	0.07	0.08	0.13	0.16	0.15	0.12	0.17	0.13	0.10	0.17	1.00	0.22	0.19	0.10
busser	0.26	0.26	0.23	0.21	0.18	0.15	0.11	0.06	0.05	0.19	0.16	0.35	0.33	0.18	0.23	0.22	1.00	0.38	0.07
govser	0.23	0.24	0.26	0.23	0.18	0.22	0.19	0.35	0.19	0.19	0.28	0.19	0.35	0.20	0.35	0.19	0.38	1.00	0.38
othser	0.22	0.26	0.18	0.36	0.28	0.46	0.36	0.58	0.51	0.19	0.44	0.12	0.38	0.26	0.38	0.10	0.07	0.38	1.00

Key concepts in PS analysis

- $PRODY_i$ (the potential income level of a product *i*) captures the extent to which a sector is associated with high or low income countries. It is a weighted average of GDP per capita of all exporting countries, with the weight of each country *c* defined by the export share for *i* in *c*:

$$PRODY_{i,t} = \sum_c \left[GDPPC_{c,t} \cdot \left(\frac{E_{c,i,t}}{\sum_{i'} E_{c,i',t}} \right) \Big/ \sum_{c'} \left(\frac{E_{c',i,t}}{\sum_{i'} E_{c',i',t}} \right) \right]$$

Key concepts in PS analysis

- $EXPY_c$ (the potential income level of the export basket of a country *c*) is a weighted average of *PRODY*s, with the weight of each sector *i* represented by its export share for *c*:

$$EXPY_{c,t} = \sum_i \frac{E_{c,i,t}}{\sum_{i'} E_{c,i',t}} \cdot PRODY_{i,t}$$

- $EXPY_c$ is a strong and robust predictor of subsequent economic growth (Hausmann et al. 2007)

Key concepts in PS analysis

- *Density* measures the capability of country *c* to become a productive producer/exporter in a sector *i* for which today its *RCA01* = 0. The definition: the share of the sum of all proximities of *i* to *i'* (*i ≠i'*) that are with sectors *i'* for which *RCA01* = 1, i.e.

$$density_{c,i|RCA01_{c,i}=0} = \sum_{i'|RCA01_{c,i'}=1} \varphi_{i,i',t} \Big/ \sum_{i'|i\neq i'} \varphi_{i,i'}$$

- Externalities across sectors: spillovers from one sector influences options for others (via technology, information, inputs, infrastructure)

Using stand-alone PS analysis to strategize

- The above PS concepts may be applied to outline a country strategy for diversification + income growth with important role for exports:
  1. Time series *RCA* data ➔ split exports into sets: classics, emerging champions, disappearing and marginals.
  2. *Density* may be used to rank products *i* with current *RCA01*=0 acc. to country capability in *i*.
  3. The products may also be ranked acc. to their PRODYs.
  4. Identifying promising products:
     - High *density* + high *PRODY* ➔ the best candidates
     - Low *density* and high *PRODY* ➔ viable in the long-run if supported by investments in inputs and/or development of other sectors.
  5. *Proximity* may be applied to a classic or emerging champion to explore whether other products, which are within close range, offer good prospects.

USING PS TO ANALYZE STRUCTURAL CHANGE IN A CGE MODEL

- Rationale 1: Enriches PS analysis – makes it possible to situate alternative PS-based strategies for structural transformation in a context that considers the role of domestic policies (taxes, public investments) and the external environment (prices, access to international markets).
- Rationale 2: Enriches CGE analysis – permits it to benefit from PS insights related to the ease/difficulty of different patterns of changes in sectoral structure.

Proximity and density in the CGE model

- Lack of *proximity* ➔ friction in the labor market via a distinction between effective and physical labor: If labor is reallocated from *i* to another sector *i'*, then the lower the proximity *i-i'*, the larger the extent to which the effective-physical labor ratio is < 1.
- The *density* of a sector *i* may influence efficiency in the same and other *i*. But need to change *density* definition so that (a)it covers all sectors: not only *i* for which *RCA01*=0; and not excluding the export performance of the sector *i* itself; and (b)provides a better measure of capability – problems with using *RCA01*:
    1. increased export share for *i* lowers *RCA* for other i';
    2. a binary variable is insensitive to change

Revealed capability instead of RCA

- Solution: In density definition, instead of *RCA01*, use *RC02* (Revealed Capability, 0 ≤ *RC02* < 2), defined as:

$$RC02_{c,i,t} = 1 + \frac{RC_{c,i,t} - 1}{RC_{c,i,t} + 1}$$

where

$$RC_{c,i,t} = \frac{\dfrac{E_{c,i,t}}{GDP_{c,t}}}{\dfrac{\sum_c E_{c,i,t}}{\sum_c GDP_{c,t}}} = \frac{\begin{bmatrix}\text{country export-GDP} \\ \text{ratio for good } i\end{bmatrix}}{\begin{bmatrix}\text{world export-GDP} \\ \text{ratio for good } i\end{bmatrix}}$$

Other problems + changes in *density* definition

- Two other problems with *density* definition:
  a. Limited to products covered by Comtrade (most goods) even though (i) capability and restructuring also is related to other sectors; and (ii) CGE analysis needs uniform treatment across sectors, especially in the labor market;
  b. It measures capability on the basis of export data; not a useful indicator of capability for relatively non-traded sectors (most of the sectors not covered by Comtrade).
- Solution: The definitions of *proximity*, *RC*, and *density* in our CGE model were expanded to non-Comtrade sectors, for these using sectoral and total GDP data, not export data.

BRINGING PS INTO MAMS

- MAMS – Maquette for MDG Simulations – is a World Bank dynamic Computable General Equilibrium (CGE) model for country-level, medium-to-long-run policy analysis, applied to around 50 countries.
- Flexible model structure (assumptions, disaggregation, years covered by analysis).
- Two versions:
  - Core: fairly standard recursive dynamic CGE model;
  - MDG: extended to include MDGs and links between education and labor market.
- This application applies the core version, extended to include PS features, to an archetype developing country database.

Details on labor/proximity and density

- In each period, for each labor category, *l*:
  1. An evolving share of the *physical* quantity is assigned to each sector *i* (based on past employment shares).
  2. In each *t*, the *physical* quantity assigned to *i* is allocated to sectors *i’* (incl. *i*).
  3. The total *effective* quantity allocated to *i’* (from any *i*) is the sum of the *physical* quantities from each *i* multiplied by the (scaled) *proximity* between *i* and *i’* (= 1 if *i* = *i’*).
  4. Complementarity slackness conditions by *l* – *i* – *i’*: [wage in *i*] ≥ [wage in *i’*]·[*proximity i* – *i’*]. Allocation $\left[QA_{l,i,i',t} > 0\right]$ permitted from *i* to *i’* if equality. (Proximity term captures shortfall of effective relative to physical labor.)
  5. In the value-added function and the first-order conditions for labor hiring, wages and quantities are expressed per unit of *effective* labor.

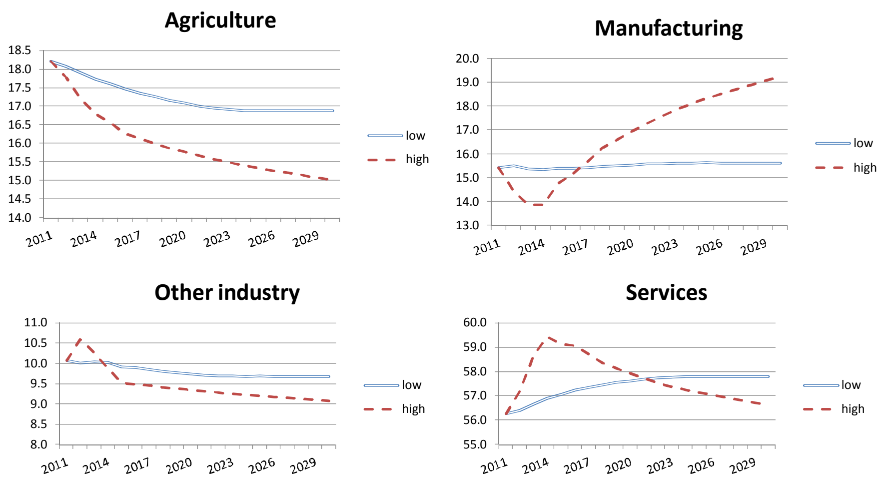
Details on labor/proximity and density

- The evolution of assignments of labor to different sectors *i* each period may be seen as reflecting learning.
- The growth in the *density* of each sector *i* in t enters a constant-elasticity function that defines the efficiency term of the value-added function of sector *i*.
- Parameters for scaling of *proximity* and elasticity of efficiency w.r.t. to *density* may be based on econometrics and validation of simulation results for initial historical period.

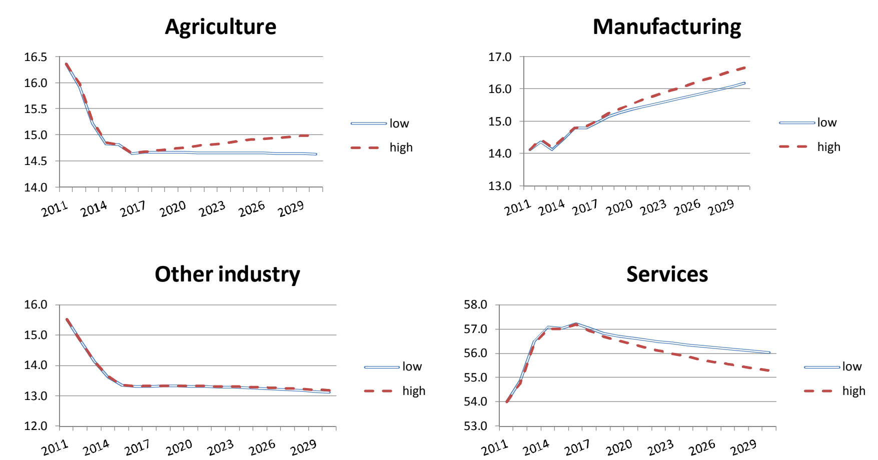
ILLUSTRATIVE AND POTENTIAL SIMULATIONS FOR ARCHETYPE COUNTRY

- Illustrative simulations: What is the impact of
  - alternative scalings of proximity parameters on the evolution of employment shares?
  - alternative values for the elasticity of sector efficiency with respect to density?

Simulated employment shares (%): low vs. high proximities (after scaling)



Simulated VA shares (%): low vs. high elasticity of efficiency with respect to density



Potential simulations

- Public infrastructure investments (broad vs. targeted) with alternative sources of financing and alternative levels of investment efficiency. What is needed for positive welfare impact?
- Replacing subsidies and indirect taxes by non-distorting direct taxes – do current policies support or discriminate against target sectors?
- Temporary subsidization of targeted sectors with alternative sources of financing – would such subsidies pay off?
- Export price responses to quality upgrading – what difference do they make to what seems like a promising strategy?
- Impact of shifts in labor force toward higher education?
- Impact of changes in world prices and trade policy?

CONCLUSIONS/ISSUES

- PS analysis offers insights that are useful for CGE modeling of long-run structural change and labor allocation.
- Indicators developed in other contexts may not be appropriate in a new CGE context.
- How disaggregated should the analysis be?
- Need for data collection/organization, econometric analysis, validation of structure/data in relation to:
  - density (alternative definitions), public infrastructure ➔ TFP growth? How/why do sectors differ?
  - proximity and friction in the labor market
  - quality upgrading and export prices

References

- Hausmann, Ricardo, Jason Hwang, and Dani Rodrik. 2007. What you export matters. Journal of Economic Growth 12 (1):1-25.
- Hausmann, Ricardo, and Bailey Klinger. 2006. Structural transformation and patterns of comparative advantage in the product space. Vol. no. 128., CID working paper. Cambridge, Mass.: Center for International Development at Harvard University.
- Hidalgo, C. A., B. Klinger, A. L. Barabasi, and R. Hausmann. 2007. The product space conditions the development of nations. Science 317 (5837):482-487.

References

- Lofgren, Hans, Martin Cicowiez, and Carolina Diaz-Bonilla, 2012. "MAMS – A Computable General Equilibrium Model for Developing Country Strategy Analysis". pp. 159–276 in Dixon, Peter B. and Dale W. Jorgenson, (Eds.), Handbook of Computable General Equilibrium Modeling. North Holland, Elsevier B.V.

- For more on MAMS, visit

  www.worldbank.org/mams

APPENDIX A. KEY CONCEPTS IN PRODUCT-SPACE ANALYSIS

PS analysis offers a data-driven evaluation of the feasibility and desirability of alternative sectoral transformation options for a country, considering its initial structure and the global record. Drawing on data for exports and, in our application, VA, it generates both standard trade indicators and a set of new indicators that are related to the *proximity* between different commodities, the income levels that are associated with the countries that export individual commodities, and the export basket of individual countries. In this section, we will review these concepts, including the extensions and adjustments that we introduce for the purposes of this study.

PS analysis was pioneered in Hausmann, Hwang, and Rodrik (2007), Hausmann and Klinger (2006), and Hidalgo et al. (2007). This analysis is based on a country's current capabilities (or initial conditions) and on international evidence that arises from patterns of comparative advantage and path dependence in the basket of goods exports of a country, implicitly underpinned by similar path dependence in the production basket of the country. PS analysis provides information about the income and diversification potential of different sector strategies, considering links between sectors. The key data source for the PS analysis at its typical, most disaggregated level, is Comtrade's global trade records (the SITC Rev. 2 Classification at the 4-digit level, covering 784 goods and 130 countries from 1980 to 2010); this information is complemented by country-level data on GDP per capita.[1]

One significant shortcoming of PS analysis is that it has typically been limited to the products (goods) covered by Comtrade, viewing these in isolation from other products.[2] Moreover, for products that are non-traded, export data cannot be used to assess the capability of a country. Furthermore, *RCA* indicators are influenced by policy-induced distortions – although, given the large number of countries involved, the resulting distortion of *proximity*-based indicators (which appear in all core PS concepts) may be quite limited – and they are not ideal proxies for capability (a concept that is distinct from comparative advantage).

Since its inception, PS analysis has been applied to a wide range of countries, including numerous studies by the World Bank that have enriched its dialogue with clients. (See the reference section for a selection of studies.)

Empirically, PS analysis involves the construction of set of export-related indicators, showing both the historical record of a country and indicative projections into the future. Some of these – including the Herfindahl index (HI), and revealed comparative advantage (*RCA*) are standard.[3]

[1] The disaggregation of PS analysis is determined by the dataset that is used. For even more finely disaggregated analysis, the Comtrade HS database may be used.

[2] In a recent pioneering analysis, Mishra, Lundstrom, and Anand (2011) address the role of service export sophistication in economic growth using the tools of PS analysis.

[3] Indicators of the technological content of exports are also frequently included in analyses of structural transformation as it is indicative of the sophistication of a country in a given commodity category. Technologically

Other indicators, derived from network theory and validated empirically, have more recently been pioneered in the context of PS analysis (Hausmann and Klinger 2006; Hidalgo and Hausmann 2009; Hidalgo et al. 2007).

More specifically, the standard indicators may be defined as follows, with the indices *r* (or *r'*), *c* (or *c'* or *c''*), and *t* referring to countries ("regions"), commodities (often referred to as products, typically limited to goods), and years, respectively:[4]

- *HI* is a time- and country-specific measure of export concentration by country, and time period:

$$HI_{r,t} = \sum_{c}\left(\frac{E_{r,c,t}}{\sum_{c'} E_{r,c',t}}\right)^2$$

  where $E_{r,c,t}$ is the value of exports for country *r* of commodity c in time *t*.

- *RCA* shows the degree of comparative advantage by country, commodity, and time. A country has an *RCA* in a commodity *c* if the following indicator has a value above unity:

$$RCA_{r,c,t} = \frac{\dfrac{E_{r,c,t}}{\sum_{c'} E_{r,c',t}}}{\dfrac{\sum_{r'} E_{r',c,t}}{\sum_{r'}\sum_{c'} E_{r',c',t}}} = \frac{\begin{bmatrix}\text{shr of commodity}\\ \text{in country exports}\end{bmatrix}}{\begin{bmatrix}\text{shr of commodity}\\ \text{in world exports}\end{bmatrix}} = \frac{\dfrac{E_{r,c,t}}{\sum_{r'} E_{r',c,t}}}{\dfrac{\sum_{i'} E_{r,c',t}}{\sum_{r'}\sum_{c'} E_{r',c',t}}} = \frac{\begin{bmatrix}\text{shr of country in}\\ \text{exports of commodity}\end{bmatrix}}{\begin{bmatrix}\text{shr of country in}\\ \text{world exports}\end{bmatrix}}$$

In the following, *RCA* refers to the above definition, i.e. a continuous variable with a lower limit of zero. The corresponding binary 0-1 variable is referred to as *RCA01* ($RCA01_{r,c,t}$ = 0 if $RCA_{r,c,t} < 1$; $RCA01_{r,c,t}$ = 1 if $RCA_{r,c,t} \geq 1$). On the basis of the evolution of their *RCA*s, exported goods may be classified as classic, emerging, disappearing, or marginal. The classics may be understood as the traditional exports of a country, i.e., commodities in which the country always has had a comparative advantage. The emerging champions are commodities in which the country did not have a comparative advantage in the past but developed it in recent years. The time periods 'past' and 'present' can be specified by the analyst. The disappearing commodities are those in which the country had a comparative advantage in the past but does not have it

sophisticated commodities tend to be associated with a high *PRODY*. For example, Lall (2000) classifies commodities into primary, resource-based, low-, medium- or high-tech; see also Lall, Weiss, and Zhang (2006).

[4] In PS analysis, the indices *c* and *i* are typically used for country and product; in MAMS and many other CGE models, *c* is used for commodities (goods or services). In order to avoid confusion and to keep notation consistent throughout this study, we switch to *r* for countries (or "regions"; this follows the example of GTAP) and use *c* for commodities. It should also be noted that MAMS makes a distinction between commodities *c* (outputs) and activities *a* (producing outputs); in this application, there is a one-to-one mapping between the two. Whether the MAMS code refers to *a* or c depends on context.

anymore, and the marginal commodities are those in which the country never has had a comparative advantage.

Indicators that more recently have been pioneered by PS analysis include:

- $PRODY_{c,t}$ indicates the potential income level of a commodity *c* in *t*. It is defined on the basis of GDP per capita (*GDPPC*) of all exporting countries, with the weight of each country defined by how important *c* is in its exports (measure by the share of *c* in the total export value of country *r*):[5]

$$PRODY_{c,t} = \sum_{r} \left[\frac{\left(\dfrac{E_{r,c,t}}{\sum_{c'} E_{r,c',t}} \right)}{\sum_{r'} \left(\dfrac{E_{r',c,t}}{\sum_{c'} E_{r',c',t}} \right)} \cdot GDPPC_{r,t} \right]$$

  Some evidence suggests that developing countries that start to produce and export a commodity over time may be able to raise the prices that they receive as quality improves, reducing the price gap vis-à-vis developed country exporters of the same commodity (Hausmann, Hwang, and Rodrik 2007, pp. 13-14).

- *EXPY* shows the potential income level of the export basket of a country. It is defined using *PRODY* with each commodity *c* weighted by its share in the exports of the country:

$$EXPY_{r,t} = \sum_{c} \frac{E_{r,c,t}}{\sum_{c'} E_{r,c',t}} \cdot PRODY_{c,t}$$

  Econometric cross-country time-series indicates that *EXPY* is a strong and robust predictor of subsequent economic growth (Hausmann, Hwang, and Rodrik 2007, p. 3).

- The *proximity* between two commodities, *c* and *c'*, in time *t,* $\varphi_{c,c',t}$ $\left(0 \le \varphi_{c,c',t} \le 1\right)$ is the key building-block for all network indicators in the PS analysis. It is derived from data on probabilities of having RCA ≥ 1 (*RCA01* = 1) simultaneously for c and c':

$$\varphi_{c,c',t} = \min\left\{ P\left(RCA01_{c,t} \middle| RCA01_{c',t} \right), P\left(RCA01_{c',t} \middle| RCA01_{c,t} \right) \right\}$$

  where *P* (the conditional probability) is computed using all countries *r* in year *t,* and where

[5] The *PRODY* concept was developed by Hausmann, Hwang and Rodrik (2007). According to Schott (2008), *PRODY* may overestimate the income potential of complex manufactured commodities such as hi tech electronics if they are exported both by relatively poor countries like China and rich Western countries. However, Schott (2008) has also noted that for simpler commodities exported by most developing countries, especially low income ones, the *PRODY* is a reasonable representation of the income potential of the commodities exported.

$$RCA01_{r,c,t} = \begin{Bmatrix} 1 \text{ if } RCA_{r,c,t} > 1 \\ 0 \text{ if } RCA_{r,c,t} \le 1 \end{Bmatrix}$$

- The *density* of a commodity *c* in which a country does *not* have an *RCA*, scaled to vary from 0 to 1, can be seen as a measure of the probability (or capability) of developing an *RCA* ≥ 1 in commodity c in the future. For each commodity, it is the ratio between (a) the sum of all proximities between that particular commodity and all commodities in which the country has an *RCA* ≥ 1; and (b) the sum of all proximities of the commodity (irrespective of whether or not the country has an *RCA* in the other commodity):

$$density_{r,c,t|RCA01_{r,c,t}=0} = \frac{\sum_{c'|c\neq c'} \varphi_{c,c',t} \cdot RCA01_{r,c',t}}{\sum_{c'|c\neq c'} \varphi_{c,c',t}}$$

In MAMS, we changed the above definition of *density*, replacing *RCA01* by an indicator we termed symmetric revealed capability (*RC02*), derived from a measure of revealed capability (*RC*):

$$RC02_{c,i,t} = 1 + \frac{RC_{c,i,t} - 1}{RC_{c,i,t} + 1}$$

and

$$RC_{c,i,t} = \frac{\dfrac{E_{c,i,t}}{GDP_{c,t}}}{\dfrac{\sum_c E_{c,i,t}}{\sum_c GDP_{c,t}}} = \frac{\begin{bmatrix} \text{country export-GDP} \\ \text{ratio for good } i \end{bmatrix}}{\begin{bmatrix} \text{world export-GDP} \\ \text{ratio for good } i \end{bmatrix}}$$

where $RC02_{c,i,t}$ = symmetric revealed capability ($0 \le RC02_{c,i,t} < 2$), $RC_{c,i,t}$ = revealed capability ($0 \le RC_{c,i,t} < \infty$). The density definition is as follows:

$$density02_{r,c,t} = \frac{\sum_{c'} \varphi_{c,c',t} \cdot RC02_{r,c',t}}{\sum_{c'|c\neq c'} \varphi_{c,c',t}}$$

Both cover all commodities (irrespective of capability level). In the implementation, the numerator is in the definition of *RC* is in real terms (at base-year prices) while the denominator is not updated (i.e., the original world ratio is kept in place). Compared to the standard *density* concept *density02* is (a) more sensitive to changes over time; this may be particularly apparent when the number of commodities is smaller than in typical PS analysis, which is the case when *density* is used in a CGE context; (b) includes the export performance of the commodity itself; (c) covers all commodities (irrespective of capability level), and (d) is less prone to generate a seeming unrealistic dampening impact on the measure of capability from a declining share in the export basket for other commodities whenever one commodity expands rapidly – no such effect would be present if other commodities maintain their own export-GDP ratios.

- Open forest (*OF*) may be seen as the option value of a country's unexploited export opportunities. It is defined on the basis of proximities to and potential incomes of all the commodities in which the country could potentially develop an *RCA*. Mathematically, it may be expressed as follows:

$$OF_{r,t} = \sum_{c} \left\{ \sum_{c'} \left[\frac{\varphi_{c,c',t}}{\sum_{c''} \varphi_{c'',c',t}} \left(1 - RCA01_{r,c',t}\right) \cdot RCA01_{r,c,t} \cdot PRODY_{c',t} \right] \right\}$$

$$= \sum_{c \mid RCA01_{r,c,t}=1} \left\{ \sum_{c' \mid rca01_{r,c',t}=0} PRODY_{c',t} \left[\frac{\varphi_{c,c',t}}{\sum_{c''} \varphi_{c'',c',t}} \right] \right\}$$

  The second formulation of the definition states that, for an individual commodity *c in which a country has an* RCA *> 1 (RCA01 = 1)*, *OF* may be defined as the potential income of every other commodity *c'* for which *RCA* < 1 *(RCA01 = 0)*, weighted by the share of *c'* in the total proximities of *c*.

When outlining an export-based country strategy for diversification and income growth in a stand-alone PS analysis, one may proceed in a stepwise manner. First, on the basis of a time series of *RCA* data, the commodities in the initial export basket of the country may be divided into four sets: classics, emerging champions, disappearing and marginals. Second, within each set, the concept of *density* can be used to rank commodities according to the country's capability to export each. Third, the commodities can be ranked according to their *PRODY*s. Commodities with a high *density* and a high *PRODY* are usually the best candidates for export expansion since high income potential for the commodity is combined with a high capability for the country to be a competitive producer of the commodity. However, in some settings, there may be a trade-off between *density* and *PRODY* which the analyst can resolve by assuming that commodities with a low *density* only are viable in the longer term when their *density* has increased, reflecting the development of an *RCA* > 1 in commodities with a high *proximity*, underpinned by investments in factors that can foster their competitive production. Finally, the concept of *proximity* can be applied to a classic or emerging champion to explore whether other commodities, which are within close range, offer prospects for export diversification. This information may be complemented by projections of world export prices and world export demand; given that the US and the EU are the leading importers of commodities exported by most developing countries, the prices of imports at the US border recorded in NAICS (North American Industry Classification), and EU import price databases may be used. World demand trends for exports may be calculated from Comtrade trade value trends. World Bank projections may be used for commodity prices.

In this study, PS indicators are generated for (a) the initial stand-alone PS analysis; and (b) use in MAMS. The stand-alone PS analysis, which only covers Comtrade commodities (most goods), is done at two levels of disaggregation: the SITC four-digit level and the disaggregation used by

MAMS. The PS indicators in MAMS follow the disaggregation of the MAMS database and cover both Comtrade and non-Comtrade commodities. For the former, the Comtrade database is used; for the latter, we use 2007 data in the GTAP 8 database. The non-Comtrade commodities consist of services and a small subset of goods (including utilities and construction). Globally and for most countries, the shares of exports in total output for these commodities are very small compared to the shares for Comtrade commodities. Given this, we used GTAP VA data for the non-Comtrade PS indicators. Technically, it is straightforward to define *RCA*, *proximity* and *density* indicators using export data for one commodity subset and VA data for another commodity subset. The economic interpretation is that, for relatively *non-tradable* commodities (covered by GTAP data), the degree of comparative (or competitive) advantage for a country is measured by how important a commodity is in the VA of that country compared to their importance in global VA. At the same time, for relatively tradable commodities (covered by Comtrade), the degree of comparative advantage for country is measured in a standard manner by the how important a commodity is in the exports of a country compared to their importance in global exports. In effect, this formulation makes it possible to consider in an integrated manner the existing patterns in terms of the development of capabilities in service sectors in parallel with goods sectors. [6] [7]

Among the above indicators, *proximity* and *density02* are most important to MAMS and its database. Drawing on *PRODY* and additional econometric analysis, MAMS will also incorporate insights about the likely evolution of export prices as a country expands exports for goods associated with countries at a higher level of GDP per capita.

REFERENCES

Hausmann, Ricardo, Jason Hwang, and Dani Rodrik. 2007. What you export matters. *Journal of Economic Growth* 12 (1):1-25.

Hausmann, Ricardo, and Bailey Klinger. 2006. *Structural transformation and patterns of comparative advantage in the product space*. Vol. no. 128., *CID working paper*. Cambridge, Mass.: Center for International Development at Harvard University.

Hidalgo, C. A., B. Klinger, A. L. Barabasi, and R. Hausmann. 2007. The product space conditions the development of nations. *Science* 317 (5837):482-487.

Lall, Sanjaya. 2000. Skills, Competitiveness and Policy in Developing Countries. In *Queen Elizabeth HouseWorking Paper*: University of Oxford.

Lall, Sanjaya, John Weiss, and Jinkang Zhang. 2006. The "sophistication" of exports: a new trade measure. *World Development (U.K.)* 34, No. 2:222-37.

[6] Note that the purpose is to assess the capability of a country to become an efficient producer in a given sector, either for exports or, when trade is less important, as a supplier of inputs or labor to sectors that are close in product space. In such an assessment, it is clearly not possible to use export data for sector without exports. However, for sectors with exports accounting for only a minor part of output (for example construction), it would not be meaningful to assess capability with reference to exports, hence the preference for using VA data.

[7] To add a further consideration, for some export-focused indicators, like *HI*, export data (globally very limited) will be used also for GTAP commodities.

Mishra, Saurabh, Susanna Lundstrom, and Rahul Anand. 2011. Service Export-Sophistication and Economic Growth. *World Bank Policy Research Working Paper*. Washington: The World Bank.
Schott, Peter. 2008. The Relative Sophistication of Chinese Exports. *Economic Policy* 23 (53):5-49.

APPENDIX B: PRODUCER PROFIT MAXIMIZATION IN A PRODUCT-SPACE SETTING

This appendix analyzes the profit-maximization problem of the producer of an activity *a* in a setting where product space *proximity* influences the profitability of the reallocation of factors between activities (sectors). The formulation of the problem is inspired by Dantzig's classical transportation problem (Dantzig 1963). A simplified version such a profit-maximization may be stated as follows:[8]

$$\text{maximize } p_a \cdot Q_a - \sum_f \sum_{a'} w_{f,a'} \cdot QFR_{f,a',a}$$

subject to

$$Q_a = \alpha_a \left[\sum_f \delta_{f,a} \cdot fp_{f,a}{}^{-\rho} \cdot QF_{f,a}{}^{-\rho} \right]^{-1/\rho} ;\; QF_{f,a} = \sum_{a'} \varphi^{scal}_{f,a',a} \cdot QFR_{f,a',a} ; \text{ and } QFR_{f,a',a} \geq 0$$

where p_a = output price; Q_a = output quantity; $w_{f,a'}$ = market wage per physical unit of factor *f* linked to in source activity *a'*; $QFR_{f,a',a}$ = the physical quantity of *f* reallocated from source activity *a'* to the current activity *a*. α_a, $\delta_{f,a}$, ρ, and $fp_{f,a}$= CES function parameters for efficiency, factor shares, exponent, and factor-specific productivity, respectively; $QF_{f,a}$ = effective employed quantity of factor *f* in activity *a*; $\varphi^{scal}_{f,a',a}$ = scaled *proximity* between any source activity *a'* and the current destination activity *a* with respect to factor *f* ($0 < \varphi^{scal}_{f,a',a} \leq 1$, with a value of 1 for the special case where *a = a'* also covered by *QFR*, i.e. "reallocation of a factor to its original activity with identical physical and effective factor quantities).[9] For factors without this factor reallocation mechanism, the *proximity* parameter is 1 for all relevant *a-a'* mappings.

The Lagrangian, *L*, for this problem may be stated as follows:

$$L = p_a \cdot Q_a - \sum_f \sum_{a'} w_{f,a'} \cdot QFR_{f,a'a} + V_a \left[\alpha_a \left[\sum_f \delta_{f,a} \cdot fp_{f,a}{}^{-\rho} \cdot QF_{f,a}{}^{-\rho} \right]^{-1/\rho} - Q_a \right]$$

[8] To simplify, this mathematical statement suppresses intermediate input use, domain controls, the time index, and the activity subscript for the exponent ρ.
[9] As noted in footnote of Appendix A, in the MAMS database, there is a one-to-one mapping between commodities, *c*, and activities, *a*. Given this, PS indicators may refer to either, depending on what is most convenient in any specific context. Since Appendix B refers to activities and not to their outputs, we here use *a*.

$$+\sum_{f} r_{f,a}\left[\sum_{a'}\varphi^{scal}_{f,a',a}\cdot QFR_{f,a',a} - QF_{f,a}\right]$$

where V_a = the shadow price of output; $r_{f,a}$ = the efficiency-unit shadow rent of factor *f* in activity *a*. The first-order conditions may be rendered as:

1. $\dfrac{\partial L}{\partial Q_a} = p_a - V_a = 0$

2. $\dfrac{\partial L}{\partial QF_{f,a}} = \left(-1/\rho\right)V_a \cdot \alpha_a\left[\sum_{f'}\delta_{f',a}\cdot fp_{f',a}{}^{-\rho}\cdot QF_{f',a}{}^{-\rho}\right]^{-1/\rho-1}\left[(-\rho)\cdot\delta_{f,a}\cdot fp_{f,a}{}^{-\rho}\cdot QF_{f,a}{}^{-\rho-1}\right] - r_{f,a} = 0$

3. $\dfrac{\partial L}{\partial QFR_{f,a',a}} = -w_{f,a'} + r_{f,a}\cdot\varphi^{scal}_{f,a',a} \le 0;\ QFR_{f,a',a} \ge 0;\ \ QFR_{f,a'a}\left(-w_{f,a'} + r_{f,a}\cdot\varphi^{scal}_{f,a',a}\right) = 0$

4. $\dfrac{\partial L}{\partial V_a} = \alpha_a\left[\sum_{f}\delta_{f,a}\cdot fp_{f,a}{}^{-\rho}\cdot QF_{f,a}{}^{-\rho}\right]^{-1/\rho} - Q_a = 0$

5. $\dfrac{\partial L}{\partial r_{f,a}} = \sum_{a'}\varphi^{scal}_{f,a',a}\cdot QFR_{f,a',a} - QF_{f,a} = 0$

Using 1 to substitute for *V* in 2, noting that $\left(-1/\rho\right)\cdot\left(-\rho\right) = 1$, and rearranging the remaining equations permits us to summarize the first-order conditions for producer profit maximization as follows:

6. $p_a\cdot\alpha_a\left[\sum_{f'}\delta_{f',a}\cdot fp_{f',a}{}^{-\rho}\cdot QF_{f',a}{}^{-\rho}\right]^{-1/\rho-1}\left[\delta_{f,a}\cdot fp_{f,a}{}^{-\rho}\cdot QF_{f,a}{}^{-\rho-1}\right] = r_{f,a}$

7. $QF_{f,a} = \sum_{a'}\varphi^{scal}_{f,a',a}\cdot QFR_{f,a',a}$

8. $w_{f,a'} \ge r_{f,a}\cdot\varphi^{scal}_{f,a',a}$

9. $QFR_{f,a',a} \ge 0$

10. $QFR_{f,a',a}\left(w_{f,a'} - r_{f,a}\cdot\varphi^{scal}_{f,a',a}\right)$

The last set of equations may be interpreted as follows:

6. The marginal value product of factor *f* in *a* (RHS) equals the shadow wage of *f* in *a* (LHS); both RHS and LHS are expressed in efficiency units.
7. Available quantity of *f* in *a* in efficiency units is the sum of quantities in physical units received from different activities *a’* multiplied by scaled *proximity*.
8. The market wage of *f* in *a’* is larger than or equal to the shadow rent of *f* in *a’* scaled by the relevant *proximity* indicator.
9. The physical quantity of *f* reallocated from any activity *a'* to *a* is non-negative.
10. If a positive quantity of *f* is reallocated from *a’* to *a*, then the market wage of *f* in source activity *a’* has to be equal to the shadow rent of *f* in destination activity *a* scaled by the

proximity indicator. (If the source wage were larger/smaller, then a marginal decrease/increase in the reallocated quantity would raise profits.) In effect, this means that, the lower the scaled *proximity*, the higher the effective-unit destination sector shadow wage that is required for reallocation to take place.

REFERENCE

Dantzig, George B. 1963. *Linear programming and extensions.* A Rand Corporation research study. Princeton, N.J.,: Princeton University Press.