Mathematical and Computational Applications

MDPI

Article

New Unidimensional Indexes for China

Gordon Bechtel

Warrington College of Business, University of Florida, P.O. Box 117150, Gainesville, FL 32611-7150, USA; bechtel@ufl.edu; Tel.: +1-352-373-4463

Academic Editor: Fazal M. Mahomed
Received: 18 December 2016; Accepted: 4 January 2017; Published: 24 January 2017

Abstract: A first principal component combines several indicators so as to maximize their internal consistency for measuring a construct. First principal components are extracted here from Swiss Economic Institute and World Bank datasets containing yearly societal indicators for China. These indicators are input to population-weighted regressions without recourse to survey sampling or probabilistic inference. The results demonstrate Chomskyan globalization and domestic credit as strong exogenous and endogenous predictors of Chinese per capita GDP. These encouraging findings, easily extendable to other nations, are brought by two new societal indexes with assured unidimensionality.

Keywords: Chomskyan globalization; domestic credit; latent principal-components and regression; per capita GDP; population-weighted indexes and regression slopes; quality assurance; ω-homogeneity

1. Introduction

Varoufakis [1] portrayed the United States of America as *The Global Minotaur*, receiving foreign capital and controlling the world economy since the Bretton Woods Conference in 1944, near the end of World War II. (http://en.wikipedia.org/wiki/Bretton_Woods_system). Central to this portrayal is John Maynard Keynes, who deplored the vindictive Versailles Treaty at the end of World War I and who saved the United States in the Great Depression, only to be dismissed at Bretton Woods by American planners. Keynes predicted that "a dollar-backed fixed exchange rate system would create other misadventures down the road […] that would […] cause […] trade imbalances to grow with, ultimately, terrible effects first upon the deficit countries and then upon everyone else" [2] (p. 10).

The most serious misadventure is described by Noam Chomsky as *technical* globalization, which denotes "a mixture of liberal and protectionist measures and many related to investor rights, not trade, all designed to serve the interests of investors, financial institutions, and other centers of concentrated state-private power" [3] (p. 35). Technical globalization has spawned dangerous trade imbalances, along with the North American Free Trade Agreement (NAFTA), the proposed Trans-Pacific Partnership (TPP), and Transatlantic Trade and Investment Partnership (TTIP). The latter proposals reach across both oceans and threaten sovereignty itself.

The pending TPP and TTIP agreements between the United States and its Pacific and European trading partners do not include the emerging BRICS economies, which stand as counterweights to these mutations in technical globalization. The capitalization and currency reserves of the BRICS's New Development Bank in Shanghai dwarf those of the World Bank and the International Monetary Fund combined. Hopefully, 21st century cooperation among these three global institutions will supplant Keynes' rejected Bretton Woods proposal: "the creation of a new world currency, a system of fixed exchange rates between this world currency and the national currencies, and a world central bank that would run the whole system" [2] (p. 14).

A second counterweight to technical globalization is developing in China's transition from manufacturing, investment, and trade toward growth that is driven by domestic services and

consumption. This economic turn from low-cost exports toward a higher standard of living for the Chinese people is paralleled by a political transition from communism to socialism. These economic and political trends in China, aided by the International Monetary Fund (IMF), World Bank, and New Development Bank, could reduce the currently dangerous trade imbalances anticipated by Keynes.

These recent forces opposing technical globalization call for quality assured economic indexes that monitor the (shortly to be) world's largest economy. The present study modifies the Swiss Economic Institute's KOF index as a proxy for *literal* globalization, which Chomsky defined as "international integration—economic, cultural, political—that serve the interests of people: real people, of flesh and blood" [3] (p. 35). KOF is an acronym for 'Konjunkturforschungsstelle', which means Economic Research Institute. The indicators in the KOF index approximate the literal meaning of globalization more closely, and are more widely used, than other measures of this construct (Samimi et al. [4]). Bechtel [5] reported that KOF globalization increased self-reported well-being in Europe. This finding, in a highly-developed region of strong unionization and corporate regulation, is extended here to the emerging and globally crucial economy of China. We find that over the period from 1992–2013, Chomskyan globalization has increased Chinese domestic credit and per capita gross domestic product (GDP). These results demonstrate the usefulness of unidimensional indexes for monitoring the sovereign performance of important economies.

Section 2 describes our indexes of globalization and domestic credit, along with the indicator of per capita GDP. Section 3 gives the unidimensional theory for the new indexes. Section 4 computes these indexes upon successive Chinese populations over the period from 1992–2013. These computations are carried out from time series supplied by the Swiss Economic Research Institute and from new World Bank software. Section 5 computes population-weighted regressions, whose slopes measure the latent effects of globalization and domestic credit on per capita GDP over the period from 1992–2013. Section 6 emphasizes the advantages of latent principal components and regression over survey regression for gauging relations between national indexes. This assessment demonstrates, without recourse to sample surveys or probabilistic inference, that Chomskyan globalization has had a positive exogenous effect upon China.

2. Constructs and Indicators

2.1. Chomskyan Globalization

The KOF Index of globalization [6] is produced by the KOF Swiss Economic Institute at ETH Zurich, who define globalization by the following indicators:

Actual flows (% of GDP) include trade, foreign direct investment, portfolio investment, and income payments to foreign nationals.
Low restrictions denote *lower* hidden import barriers, mean tariff rate, taxes on international trade, and capital account restrictions.
Personal contact includes telephone traffic, transfers, foreign population, and international letters.
Information flows refer to internet users, television, and trade in newspapers.
Cultural proximity is made up of trade in books and numbers of McDonald's and Ikea.
Political globalization consists of embassies in country, membership in international organizations, participation in U.N. Security Council missions, and international treaties.

The Swiss Economic Research Institute regards *actual flows* and *low restrictions* as economic globalization and *personal contact*, *information flows*, and *cultural proximity* as social globalization. We use KOF's six indicators as proxies for Chomsky's literal globalization, defined as 'international integration—economic, cultural, political' [3] (p. 35).

2.2. Domestic Credit

The International Monetary Fund (IMF) observed that the declining surplus in China is due to (a) the collapse of global demand caused by the financial crisis of 2008 and (b) China's steep increase

in investment "to prop up domestic demand and offset the large shock emanating from the […] collapse in external demand" [7]. In generating this offset, the IMF noted that "China already has a head start compared with other high-growth economies with regard to financial deepening. Bank deposits and credit as shares of GDP are higher in China than in many other economies" [7]. Domestic credit is measured by two indicators monitored by the World Bank:

> **Domestic credit *to* the private sector (% of GDP)** refers to financial resources provided to the private sector by financial corporations. The financial corporations include monetary authorities and deposit money banks.
> **Domestic credit *provided by* the financial sector (% of GDP)** includes all credit to various sectors. The financial sector includes monetary authorities and deposit money banks.
>
> Chinese domestic credit provides a huge potential for kinetic household demand and the business investment that supplies it.

2.3. Per Capita GDP

The IMF [7] notes that Chinese government policy is now "designed to accelerate the transformation of the Chinese economic model, improve livelihoods, and raise domestic consumption. [...] If successful, this would ultimately prove to […] contribute significantly to strong, sustained, and balanced global growth." The importance of sustained GDP growth has also been emphasized by the Leeds UK Steady State Economy Conference [8], the United Nations Division for Sustainable Development [9], and the Annual Forum of The Progressive Economy Initiative [10]. This economic consensus points to *per capita GDP* as a policy objective, rather than the unbridled *GDP growth* of technical globalization (cf. [3]):

> **GDP per capita** is the gross domestic product divided by the midyear population. Data are in constant 2010 U.S. dollars.

This World Bank index of sovereign performance (http://beta.data.worldbank.org) captures China's turn from low-cost exports toward domestic services and consumption.

3. Unidimensional Index Theory

3.1. Indicator Weighting

Unidimensional indexes are obtained here by principal components analysis [11] (pp. 536–544), which finds optimal indicator weights that maximize *internal* index consistency. This approach differs from the KOF index and national index theory [12], which weight indicators a priori. It also differs from recent Chinese indexes, which weight their indicators to maximize the prediction of *external* criteria such as growth.

The theory in Sections 3.2 and 3.3 is written for the six globalization indicators in Section 2.1. Section 3.4 is written for globalization, domestic credit, and per capita GDP, whose indicators are described in Sections 2.1, 2.2, and 2.3.

3.2. Latent Principal Components

Existential Axiom 1. *In year t, China's globalization indicator G_{tj} (j = 1, …, 6) induces a latent idiosyncratic impact G_{tij} on individual i = 1, …, N_t, where N_t is China's population size in year t = 1992, …, 2013.*

Homogeneity Axiom 2. *The within-year covariance matrix of the vector (G_{ti1} G_{ti2} … G_{ti6}) over i = 1, …, N_t and t = 1992, …, 2013 is ω times its between-year covariance matrix, where ω is an unknown positive scalar.*

Corollary 1. *The total covariance matrix of the vector (G_{ti1} G_{ti2} … G_{ti6}) is ($1 + \omega$) times its between-year covariance matrix.*

Next, we linearly transform G_{tij} to Z_{tij} for j = 1, …, 6, such that the distribution of each Z_{tij} has a mean of zero and a variance one over i = 1, …, N_t and t = 1992, …, 2013. This standardizes the different units of G_{ti1}, G_{ti2}, G_{ti3}, G_{ti4}, G_{ti5}, and G_{ti6}.

Lemma 1. *The within-year covariance matrix of the vector (Z_{ti1} Z_{ti2} … Z_{ti6}) over i = 1, …, N_t and t = 1992, …, 2013 is also ω times its between-year covariance matrix, i.e. the multiplicative relation of within-to-between-year covariance matrices is preserved under distinct linear transformations of the six latent globalization impacts.*

Corollary 2. $\Sigma_t\Sigma_i Z_{tij}^2 = (1 + \omega)\Sigma_t N_t Z_{t\bullet j}^2$ *and* $\Sigma_t\Sigma_i Z_{tij}Z_{tik} = (1 + \omega)\Sigma_t N_t Z_{t\bullet j}Z_{t\bullet k}$, *where the subscripted dots denote averages over i = 1, …, N_t for t = 1992, …, 2013.*

Lemma 2. *The latent correlation between Z_{tij} and Z_{tik} may be written as*

$$r_{jk} = \Sigma_t N_t Z_{t\bullet j}Z_{t\bullet k} / (\Sigma_t N_t Z_{t\bullet j}^2)^{1/2}(\Sigma_t N_t Z_{t\bullet k}^2)^{1/2}.$$

Corollary 3. *Individual ti's latent score on the first principal component* $\mathbf{Z}$ *of matrix (r_{jk}) is*

$$Z_{ti} = a_1 Z_{ti1} + a_2 Z_{ti2} + \ldots + a_6 Z_{ti6},$$

where the vector (a_1 a_2 … a_6) is the first eigenvector of (r_{jk}) [11] (pp. 536-537).

Lemma 3. *If $a_1^2 + a_2^2 + \ldots + a_6^2$ is the reciprocal of the first eigenvalue of (r_{jk}), then the first latent principal component* $\mathbf{Z}$ *is also standardized with mean zero and variance one over i = 1, …, N_t and t = 1992, …, 2013.*

Corollary 4. $\mathbf{Z} = \mathbf{Z}_b + \mathbf{Z}_w$, *where* $\mathbf{Z}_b$, *with variance $(1 + \omega)^{-1}$, replicates the yearly mean*

$$Z_{t\bullet} = \Sigma_i Z_{ti}/N_t = a_1 Z_{t\bullet 1} + a_2 Z_{t\bullet 2} + \ldots + a_6 Z_{t\bullet 6}$$

over i = 1, …, N_t for t = 1992, …, 2013.

The within-year vector $Z_w = Z - Z_b$, with variance $\omega(1 + \omega)^{-1}$, contains unknowable deviations $Z_{ti} - Z_{t\bullet}$ over i = 1, …, N_t and t = 1992, …,2013. However, the between-year vector Z_b is knowable up to an unknown multiplier. We now express this multiple of Z_b.

3.3. Manifest Principal Components

Definition. $\mathcal{Z}_{tj}$ *is the N_t-weighted standardized value of G_{tj} in axiom 1.*

Lemma 4. *Due to Corollary 2,* $\mathcal{Z}_{tj} = (1 + \omega)^{1/2} Z_{t\bullet j}$.

Lemma 5. *Due to Lemmas 2 and 4, the latent correlation r_{jk} in lemma 2 between Z_{tij} and Z_{tik} may be expressed as the manifest correlation*

$$\mathcal{R}_{jk} = \Sigma_t N_t \mathcal{Z}_{tj}\mathcal{Z}_{tk} / (\Sigma_t N_t \mathcal{Z}_{tj}^2)^{1/2}(\Sigma_t N_t \mathcal{Z}_{tk}^2)^{1/2}.$$

Corollary 5. *The eigenvalues and eigenvectors of correlation matrices (r_{jk}) and ($\mathcal{R}_{jk}$) are identical.*

Corollary 6. *Corollary 4 and Lemma 4 give China's yearly globalization index*

$$\mathcal{G}_t = (1 + \omega)^{1/2} Z_{t\cdot} = a_1 \mathcal{Z}_{t1} + a_2 \mathcal{Z}_{t2} + \ldots + a_6 \mathcal{Z}_{t6} \quad \text{for } t = 1992, \ldots, 2013.$$

Corollary 7. *An N_t-weighted principal components analysis of the matrix $(\mathcal{Z}_{tj})$ has first principal component $\boldsymbol{\mathcal{G}} = (1 + \omega)^{1/2} \mathbf{Z}_b$, which replicates $\mathcal{G}_t$ over $i = 1, \ldots, N_t$ for $t = 1992, \ldots, 2013$.*

Note that first manifest principal component $\boldsymbol{\mathcal{G}}$ is standardized with mean zero and variance of one. Also, manifest index $\mathcal{G}_t$ is the product of two unobservable factors; the multiplier $(1 + \omega)^{1/2}$ and the latent index $Z_{t\cdot}$.

3.4. Latent Regression

Our second index consists of the two indicators in Section 2.2. Therefore, in Sections 3.2 and 3.3 we let $j = 1, 2$ and overwrite 'globalization' by 'domestic credit', $\mathcal{G}$ by $\mathcal{D}$, G by D, and Z by Y. This allows the interpretation of the resulting index $\mathcal{D}_t = (1 + \omega)^{1/2} Y_{t\cdot}$ as the product of multiplier $(1 + \omega)^{1/2}$ and the latent mean impact $Y_{t\cdot}$ of domestic credit for year t.

Our final index is the single indicator in Section 2.3. Thus, we set $j = 1$ in Sections 3.2 and 3.3 and overwrite 'globalization' by 'per capita GDP', $\mathcal{G}$ by $\mathcal{W}$, G by W, and Z by X. This permits us to interpret China's manifest index $\mathcal{W}_t = (1 + \omega)^{1/2} X_{t\cdot}$ as the product of unknown factors $(1 + \omega)^{1/2}$ and the latent per capita GDP index $X_{t\cdot}$.

Principal components $\boldsymbol{\mathcal{G}}$ and $\boldsymbol{\mathcal{D}}$, and indicator $\boldsymbol{\mathcal{W}}$, replicate $\mathcal{G}_t$, $\mathcal{D}_t$, and $\mathcal{W}_t$ over $i = 1, \ldots, N_t$ for $t = 1992, \ldots, 2013$. These replicated manifest indexes are multiples of China's latent indexes of globalization $Z_{t\cdot}$, domestic credit $Y_{t\cdot}$, and per capita GDP $X_{t\cdot}$ in year t. The observable sums-of-products matrix of $\boldsymbol{\mathcal{G}}$, $\boldsymbol{\mathcal{D}}$, and $\boldsymbol{\mathcal{W}}$,

$$\begin{pmatrix} \boldsymbol{\mathcal{G}}^T\boldsymbol{\mathcal{G}} & \boldsymbol{\mathcal{G}}^T\boldsymbol{\mathcal{D}} & \boldsymbol{\mathcal{G}}^T\boldsymbol{\mathcal{W}} \\ \boldsymbol{\mathcal{D}}^T\boldsymbol{\mathcal{G}} & \boldsymbol{\mathcal{D}}^T\boldsymbol{\mathcal{D}} & \boldsymbol{\mathcal{D}}^T\boldsymbol{\mathcal{W}} \\ \boldsymbol{\mathcal{W}}^T\boldsymbol{\mathcal{G}} & \boldsymbol{\mathcal{W}}^T\boldsymbol{\mathcal{D}} & \boldsymbol{\mathcal{W}}^T\boldsymbol{\mathcal{W}} \end{pmatrix} \quad (1)$$

enables latent regressions involving globalization, domestic credit, and per capita GDP.

Homogeneity Axiom 3. $\mathbf{Z}_w^T\mathbf{Y}_w/\mathbf{Z}_b^T\mathbf{Y}_b = \mathbf{Z}_w^T\mathbf{X}_w/\mathbf{Z}_b^T\mathbf{X}_b = \mathbf{Y}_w^T\mathbf{X}_w/\mathbf{Y}_b^T\mathbf{X}_b = \varpi$.

Bivariate Lemma 6. *The manifest regression slope $\boldsymbol{\mathcal{G}}^T\boldsymbol{\mathcal{D}}/\boldsymbol{\mathcal{G}}^T\boldsymbol{\mathcal{G}}$ of domestic credit on globalization equals the latent slope $\mathbf{Z}^T\mathbf{Y}/\mathbf{Z}^T\mathbf{Z}$, i.e.* $\mathbf{Z}^T\mathbf{Y}/\mathbf{Z}^T\mathbf{Z} = (\mathbf{Z}_b^T\mathbf{Y}_b + \mathbf{Z}_w^T\mathbf{Y}_w)/(\mathbf{Z}_b^T\mathbf{Z}_b + \mathbf{Z}_w^T\mathbf{Z}_w) = (\mathbf{Z}_b^T\mathbf{Y}_b + \varpi\mathbf{Z}_b^T\mathbf{Y}_b)/(\mathbf{Z}_b^T\mathbf{Z}_b + \varpi\mathbf{Z}_b^T\mathbf{Z}_b)$ $= (1 + \varpi)\mathbf{Z}_b^T\mathbf{Y}_b/(1 + \varpi)\mathbf{Z}_b^T\mathbf{Z}_b = (1 + \varpi)^{1/2}\mathbf{Z}_b^T(1 + \varpi)^{1/2}\mathbf{Y}_b/(1 + \varpi)^{1/2}\mathbf{Z}_b^T(1 + \varpi)^{1/2}\mathbf{Z}_b = \boldsymbol{\mathcal{G}}^T\boldsymbol{\mathcal{D}}/\boldsymbol{\mathcal{G}}^T\boldsymbol{\mathcal{G}}$.

The matrix generalization of lemma 6 is multivariate lemma 7.

Multivariate Lemma 7. *The manifest and latent regression slopes of per capita GDP on globalization and domestic credit are equal, i.e.*

$$\begin{pmatrix} \mathbf{Z}^T\mathbf{Z} & \mathbf{Z}^T\mathbf{Y} \\ \mathbf{Y}^T\mathbf{Z} & \mathbf{Y}^T\mathbf{Y} \end{pmatrix}^{-1} \begin{pmatrix} \mathbf{Z}^T\mathbf{X} \\ \mathbf{Y}^T\mathbf{X} \end{pmatrix} = \begin{pmatrix} \boldsymbol{\mathcal{G}}^T\boldsymbol{\mathcal{G}} & \boldsymbol{\mathcal{G}}^T\boldsymbol{\mathcal{D}} \\ \boldsymbol{\mathcal{D}}^T\boldsymbol{\mathcal{G}} & \boldsymbol{\mathcal{D}}^T\boldsymbol{\mathcal{D}} \end{pmatrix}^{-1} \begin{pmatrix} \boldsymbol{\mathcal{G}}^T\boldsymbol{\mathcal{W}} \\ \boldsymbol{\mathcal{D}}^T\boldsymbol{\mathcal{W}} \end{pmatrix},$$

where the manifest 2 × 2 and 2 × 1 matrices are submatrices of the matrix in *(1).*

The slopes in lemmas 6 and 7 derive from multi-year population regressions of 28 billion latent impacts over $i = 1, \ldots, N_t$ and $t = 1992, \ldots, 2013$ (cf. Axiom 1). These slopes are not subject to sampling variation like regression slopes estimated from survey samples. Hence, we are unencumbered by probabilistic inference, estimation, or hypothesis testing (cf. Section 6).

3.5. Sufficiency (but not Necessity) of ϖ-Homogeneity

Axioms 1 and 2 imply the manifest indexes $\mathcal{G}_t$ for t = 1992, …, 2013. However, population-weighted principal components analysis produces identical manifest indexes in the presence of other latent multi-year distributions that are not ϖ-homogeneous. Thus, ϖ-homogeneity is a sufficient, but not necessary, condition for the observation of manifest indexes $\mathcal{G}_t$ for t = 1992, …, 2013.

Axioms 1–3, which posit ⊚-homogeneity within and between indexes, imply the manifest regression slopes in lemmas 6 and 7. Again, population-weighted regression analysis produces identical manifest slopes from other latent multi-year distributions that are not ⊚-homogeneous. Therefore, ⊚-homogeneity is also sufficient, but not necessary, for the observation of manifest slopes identical to those in Lemmas 6 and 7.

The sufficiency (but not necessity) of ϖ-homogeneity signifies a broader class of latent multi-year distributions supporting the manifest indexes and slopes in Sections 3.2–3.4. Thus, the class of ϖ-homogeneous distributions is a canonical subset of the class of latent distributions that give identical population-weighted indexes and regression slopes.

4. Computation of Population-Weighted Indexes

4.1. Globalization Index

The six indicators in the following two Stata [13] commands are described in Section 2.1:

```
factor flows lowrestrict contact infoflows culturprox politglobal [fweight = popt], pcf
predict global, norotated
```
(2)

The option pcf denotes principal components factorization. The optional qualifier [fweight = popt] gives an N_t-weighted first principal component 'global' that replicates indexes $\mathcal{G}_t$ over $i = 1, \ldots, N_t$ for $t = 1992, \ldots, 2013$. The Stata commands in (2) operate on a 22 × 6 spreadsheet.

Table 1 exhibits the eigenvalues found by our analysis of Chomskyan globalization. The unidimensionality obtained from the first principal component is demonstrated by the dominance of the first eigenvalue over the remaining five eigenvalues. This implies that $\mathcal{G}$ extracts the preponderance of information contained in these six indicators, i.e., that these indicators are internally consistent in measuring a single construct. This first principal component accounts for 4.897/6 = 0.816 of the variance in these KOF indicators. Thus, our new index of Chomskyan globalization is quality assured for China.

Table 1. Eigenvalues for globalization.

Principal Components	1	2	3	4	5	6
Eigenvalues	4.897	0.445	0.393	0.194	0.061	0.010

Notes: the proportion of indicator variance accounted for by the first principal component is 4.897/6 = 0.816.

Table 2 details the internal index consistency demonstrated in Table 1. Each loading in the loadings eigenvector is the Pearson-product-moment correlation of that indicator with the first principal component over $i = 1, \ldots, N_t$ and $t = 1992, \ldots, 2013$. Due to Corollary 5, this manifest loading also equals the latent correlation between Z_{tij} and $Z_{ti,}$ i.e., between the individual latent impacts of indicator j (= 1, … ,6) and those of the first principle component over $i = 1, \ldots, N_t$ and $t = 1992, \ldots, 2013$. Finally, each coefficient in the scoring eigenvector is the weight of that indicator in individually scoring this first principal component (cf. Corollary 3). The sum of the squared scoring coefficients is the reciprocal of the first eigenvalue in Table 1 (cf. Lemma 3).

Table 2. Components of the first principal component for globalization.

Indicators	Flows	Lowrestrict	Contact	Infoflows	Culturprox	Politglobal
Loadings	0.841	0.790	0.946	0.968	0.896	0.964
Scoring coefficients	0.172	0.161	0.193	0.198	0.183	0.197

Notes: the sum of the squared loadings equals the first eigenvalue in Table 1. The scoring eigenvector is a scalar multiple of the loadings eigenvector. The sum of the squared scoring coefficients is the reciprocal of the first eigenvalue in Table 1.

4.2. Domestic Credit Index

Section 2.2 describes the indicators in our two commands for domestic credit:

$$\begin{array}{c}\text{factor privatecredit financialcredit [fweight = popt], pcf}\\ \text{predict domesticcredit, norotated}\end{array} \quad (3)$$

The N_t-weighted first principal component 'domesticcredit' replicates indexes $\mathcal{D}_t$ over $i = 1, \ldots, N_t$ for $t = 1992, \ldots, 2013$. The Stata commands in (3) operate on a 22 × 2 spreadsheet.

Tables 3 and 4 exhibit the eigenvalues and eigenvectors for domestic credit. The strong unidimensionality of the first principal component $\mathcal{D}$ is seen by the large difference between the two eigenvalues. The first principal component accounts for 1.974/2 = 0.987 of the variance in the two indicators, whose loadings on this factor are also extremely high. This degree of internal consistency quality-assures our new credit index for China.

Table 3. Eigenvalues for domestic credit.

Principal Components	1	2
Eigenvalues	1.974	0.026

Notes: the proportion of indicator variance accounted for by the first principal component is 1.974/2 = 0.987.

Table 4. Eigenvectors of the first principal component of domestic credit.

Indicators	Private Sector Credit	Financial Sector Credit
Loadings	0.993	0.993
Scoring coefficients	0.503	0.503

Notes: the sum of the squared loadings equals the first eigenvalue in Table 3. The scoring eigenvector is a scalar multiple of the loadings eigenvector. The sum of the squared scoring coefficients is the reciprocal of the first eigenvalue in Table 3.

5. Computation of Population-Weighted Regression Slopes

5.1. The Setup for Regression

Our latent bi-variate and multi-variate regressions run over individuals $i = 1, \ldots, N_t$ in years $t = 1992, \ldots, 2013$, encompassing 28 billion personal impacts. The regression slopes read out from Lemmas 6 and 7 are computed from manifest principal components $\mathcal{G}$ and $\mathcal{D}$ and indicator $\mathcal{W}$. These three vectors carry the replicated yearly indexes $\mathcal{G}_t$, $\mathcal{D}_t$, and $\mathcal{W}_t$. The standardizations of the two manifest principal components, and the single indicator, to mean zero and variance one, enables a comparison of the slopes of $\mathcal{W}$ on $\mathcal{G}$ and $\mathcal{D}$.

5.2. Bivariate Regression of Principal Components

The strong exogenous effect of Chomskyan globalization on domestic credit is 0.806 with $R^2 = 0.650$. This regression slope, derived in Lemma 6, was computed from Stata [13] Command

$$\text{regress domesticcredit global [fweight = popt],} \quad (4)$$

where popt is China's population in year t = 1992, …, 2013. The optional qualifier [fweight = popt] gives the Nt-weighted slope $\mathcal{G}^T\mathcal{D}/\mathcal{G}^T\mathcal{G}$. This manifest slope equals the latent slope $\boldsymbol{Z}^T\boldsymbol{Y}/\boldsymbol{Z}^T\boldsymbol{Z}$ when the within-year covariance matrix of $\boldsymbol{Y}$ and $\boldsymbol{Z}$ is ω times their between-year covariance matrix.

5.3. Multivariate Regression of per Capita GDP on Principal Components

Table 5 exhibits the slopes and R^2 for the regression of per capita GDP on globalization and domestic credit. The multi-year population slopes in Table 5 were computed from Stata command

regress gdp global domesticcredit [fweight = popt], beta (5)

which runs an N_t-weighted regression over t = 1992, …, 2013. This regression gives manifest slopes that equal their latent counterparts when the within-year covariance matrix of $\boldsymbol{Z}$, $\boldsymbol{Y}$, and $\boldsymbol{X}$ is ω times their between-year covariance matrix (cf. Lemma 7). The beta option standardizes the GDP indicator to mean zero and variance one, and maintains this prior standardization of the principal components for domestic credit and globalization. This option allows a comparison of the two regression slopes delivered by Command (5).

Regressions (4) and (5) demonstrate that Chomskyan globalization exogenously drives domestic credit which, in turn, endogenously drives per capita GDP. Table 5 shows that Chomskyan globalization also exerts a direct exogenous impact on per capita GDP. A high proportion (R^2 = .689) of the variance in $\mathcal{W}$ is accounted for by $\mathcal{G}$ and $\mathcal{D}$. These results extend Chomskyan globalization's beneficial impacts on domestic well-being from Europe [5] and the G20 nations [12] to China.

Table 5. Regression effects on per capita GDP.

Chomskyan Globalization	Domestic Credit	R^2
0.407	0.466	0.689

Notes: the two slopes were computed from Lemma 7 and Stata Command (5).

6. Future Directions

Unidimensional index theory overrides "the central dogma of statistical inference, that there is a component of randomness in data" [14] (p. 9). "Neither denying nor quantifying uncertainty, we simply ignore it" [12] (p. 9).

The axiomatic approach to sequential human populations in Section 3 brings compelling advantages to the data science of human populations. Probabilistic inference is replaced by parameter computation and random variables give way to real variables. This suggests further "statistical thinking and new foundational frameworks" that help sort out "the many philosophical issues data science presents" [15] (pp. 4,5). This call has been echoed by the American National Science Foundation, which has "released a revised version of the solicitation Critical Techniques and Technologies for Advancing Foundations and Applications of Big Data Science" [16] (p. 6).

6.1. Advantages of Latent Principal Components and Regression over Sample Surveys

Data definition. Pfeffermann observed that "The use of big data does not require a sampling frame, questionnaires, interviews, and all the other ingredients underlying survey samples. […] this should be the ultimate target of every country—having sufficiently accurate administrative records so that no population censuses will be needed" [17] (pp. 433, 455). Horrigan also views Big Data as nonsampled data "from electronic sources whose primary purpose is something other than statistical inference […] this type of Big Data typically comprises the universe and, by definition, can represent (nearly) the entire population" [18] (pp. 25–26). As examples of nonsampled universe files, Horrigan mentions daily price indexes, point-of-sale retail databases, universe data on hospitals, and corporate data.

The present paper exploits the electronic files of the World Bank and the Swiss Economic Research Institute. Principal-component indexes, directly computed from established macro indicators, are more plausible and palpable than constructs hypothesized to be measured by survey

samples of human populations. Latent principle-component scores (cf. Corollaries 3 and 4) are also practical alternatives to questionnaire scores that require the protection of individual privacy. In most countries, self-reported consumption, received services, and well-being are not even available.

Data collection. For almost half a century the interrogation of individuals in survey samples has been shadowed by skepticism about the incremental benefits of subjective indicators over and above objective indicators already in use [19–22]. The problems associated with subjective measurement point up the pitfalls of survey sampling, which may not be needed in the first place (cf. [23,24]).

Population-weighted indexes and regressions evade the host of long-standing, and now acute, issues daunting micro-data collection. First, sampled micro data are beset by the unresolved competition between randomization-based and model-based regression [25–27]. Second, both types of regression face problems of measurement error [28], sampling error [29,30], unit nonresponse [31], missing data [32], and variance estimation [33,34]. Unit nonresponse alone threatens the entire survey industry due to public unwillingness to answer mail, telephone, internet, or face-to-face questions [35]. The host of problems associated with survey measurement and process quality were discussed and illustrated almost twenty years ago in the volume edited by Lyberg et al. [23]. Today Bradburn [24] (p. 94) laments that "the challenges confronting the survey researcher are dominated by the difficulty in locating sample persons and getting them to respond at all." This reluctance to respond has spawned "Suspicious Supervisors and Suspect surveys […] preventing data fabrication is a widely understood challenge among survey professionals that is seldom discussed publicly" [36].

Data analytics. Here latent regression slopes over 28 billion values are read out in Lemmas 6 and 7 and computed by Stata Commands (4) and (5). This computation is justified by the theory in Section 3, which derives latent Chinese indexes as first moments of unobservable yearly distributions.

Lemma 6, implemented by Command (4), reveals the effect of globalization on domestic credit. The slope 0.806 of $\mathcal{D}$ on $\mathcal{G}$, with $R^2 = 0.650$, shows that Chomskyan globalization has exogenously increased Chinese domestic credit over the period from 1992–2013. Table 5 then shows that the endogenous effect of domestic credit *and* the exogenous impact of Chomskyan globalization account for almost 70% of the variance in Chinese per capita GDP.

6.2. Conclusions

Unidimensional index theory weights economic indicators so as to maximize the internal consistency of their societal indexes. This approach differs from indexing that weights indicators a priori or that maximizes the prediction of external criteria such as growth.

The present results demonstrate Chomskyan globalization and domestic credit as strong exogenous and endogenous predictors of per capita GDP in China. These encouraging findings, easily extendable to other nations, are brought by two new macro indexes with assured unidimensionality. In the future, other economic indicators, discovered from Big Data [37–40], can be input to Commands (2) and (3) to quality assure other indexes. The discovery of new unidimensional indexes, facilitated by cooperating governments, the United Nations, the World Bank, the International Monetary Fund, and the New Development Bank in Shanghai, will help data science tackle trade imbalances and upgrade societal well-being 'after the Minotaur' [41].

Acknowledgments: The author thanks Bethany Bechtel for her penetrating review of this work, physical analogues of unobservable human populations, and emphasis on monitoring these populations over time. The author also thanks the reviewers, whose guidance and critiques have improved the substance and presentation of this paper.

Conflicts of Interest: The author declares no conflicts of interest.

References

1. Varoufakis, Y. *The Global Minotaur: America, the True Origins of the Financial Crisis and the Future of the World Economy*; Zed Books: London, UK, 2011.
2. Varoufakis, Y. *And the Weak Suffer What They Must? Europe's Crisis and America's Economic Future*; Nation Books: New York, NY, USA, 2016.

3. Chomsky, N. *Hopes and Prospects*; Haymarket Books: Chicago, IL, USA, 2010.
4. Samimi, P.; Lim, G.C.; Buang, A. A critical review of synthetic globalization indexes. *Int. J. Fundam. Psychol. Soc. Sci.* **2012**, *2*, 28–31.
5. Bechtel, G.G. Does globalization mitigate income inequality? *J. Data Sci.* **2014**, *12*, 197–215.
6. Dreher, A. Does globalization affect growth? Evidence from a new index of globalization. *Appl. Econ.* **2006**, *38*, 1091–1110.
7. Singh, A.; Nabar, M.; N'Diaye, P. (Eds.) *China's Economy in Transition: From External to Internal Rebalancing*; International Monetary Fund: Washington, DC, USA, 2013.
8. O'Neill, D.W.; Dietz, R.; Jones, N. (Eds.) *Enough is Enough: Ideas for a Sustainable Economy in a World of Finite Resources*; The Report of the Steady State Economy Conference; Center for the Advancement of the Steady State Economy and Economic Justice for All: Leeds, UK, 2010.
9. Costanza, R.; Alperovitz, G.; Daley, H.; Farley, J.; Franco, C.; Jackson, T.; Kubiszewski, I.; Schor, J.; Victor, P. *Building a Sustainable and Desirable Economy-in-Society-in-Nature*; United Nations Division for Sustainable Development: New York, NY, USA, 2012.
10. Progressive Economy. *A Call for Change: From Crisis to a New Egalitarian Ideal for Europe*; The Progressive Economy Initiative: Brussels, Belgium, 2014.
11. Johnston, J. *Econometric Methods*, 3rd ed.; McGraw-Hill: New York, NY, USA, 1984.
12. Bechtel, G.G. National index theory: Chomskyan globalization and well-being. *Comput. Appl. Math. J.* **2016**, *2*, 6–11.
13. StataCorp. *Stata Statistical Software*; Release 12; StataCorp LP: College Station, TX, USA, 2011.
14. Van Dyk, D.; Fuentes, M.; Jordan, M.; Newton, M.; Ray, B.; Lang, D.; Wickham, H. ASA Statement on the Role of Statistics in Data Science. *Amstat News*, October 2015, p. 9.
15. Davidian, M. Aren't *We* Data Science? *Amstat News*, July 2013, pp. 3–5.
16. Vogelius, M.; Kannan, N.; Huo, X. NSF Big Data Funding Opportunity for the Statistics Community. *Amstat News*, April 2015; p. 6.
17. Pfeffermann, D. Methodological issues and challenges in the production of official statistics, 24th Annual Morris Hansen Lecture. *J. Surv. Stat. Methodol.* **2015**, *3*, 425–467.
18. Horrigan, M.W. Big Data: A Perspective from the BLS. *Amstat News*, January 2013, pp. 25–27.
19. Shapiro, H. The Index of Consumer Sentiment and economic forecasting: A reappraisal. In *Human Behavior in Economic Affairs*; Strumpel, B., Morgan, J.N., Zahn, E., Eds.; Jossey-Bass: San Francisco, CA, USA, 1972; Chapter 18.
20. Didow, N.M.; Perreault, W.D.; Williamson, N.C. A cross-sectional optimal scaling analysis of the index of consumer sentiment. *J. Consum. Res.* **1983**, *10*, 339–347.
21. Vanden Abeele, P.V. The index of consumer sentiment: Predictability and predictive power in the EEC. *J. Econ. Psychol.* **1983**, *3*, 1–17.
22. Bram, J.; Ludvigson, S. *Does Consumer Confidence Forecast House-Hold Expenditure? A Sentiment Index Horse Race*; Federal Reserve Bank of New York: New York, NY, USA, 1998; pp. 59–78.
23. Lyberg, L.; Biemer, P.; Collins, M.; de Leeuw, E.; Dippo, C.; Schwarz, N.; Trewin, D. (Eds.) *Survey Measurement and Process Quality*; Wiley and Sons: New York, NY, USA, 1997.
24. Bradburn, N.M. Surveys as social interactions. *J. Surv. Stat. Methodol.* **2016**, *4*, 94–109.
25. Chaudhuri, A.; Stenger, H. *Survey Sampling: Theory and Methods*, 2nd ed.; CRC Press: Boca Raton, FL, USA, 2005.
26. Brewer, K.; Gregoire, T.G. Introduction to survey sampling. In *Handbook of Statistics (Sample Surveys: Design, Methods and Applications)*; Pfeffermann, D., Rao, C.R., Eds.; Elsevier: Amsterdam, The Netherlands, 2009; Volume 29A, Chapter 1.
27. Opsomer, J.D. Introduction to Part 4: Alternative approaches to inference from survey data. In *Handbook of Statistics (Sample Surveys: Inference and Analysis)*; Pfeffermann, D., Rao, C.R., Eds.; Elsevier: Amsterdam, The Netherlands, 2009; Volume 29B, pp. 3–9.
28. Bound, J.; Brown, C.; Mathiowetz, N. Measurement error in survey data. In *Handbook of Econometrics*; Heckman, J.; Leamer, E., Eds.; Elsevier Science: Amsterdam, The Netherlands, 2001; Volume 5, Chapter 59.
29. Thompson, M.E. *Theory of Sample Surveys*; Chapman and Hall: London, UK, 1997.
30. Lohr, S.L. *Sampling: Design and Analysis*, 2nd ed.; Brooks/Cole: Boston, MA, USA, 2010.
31. Brick, J.; Montaquila, J.M. Nonresponse and weighting. In *Handbook of Statistics (Sample Surveys: Design, Methods and Applications)*; Pfeffermann, D., Rao, C.R., Eds.; Elsevier: Amsterdam, The Netherlands, 2009; Volume 29A, Chapter 8.

32. Little, R.J.A.; Rubin, D.B. *Statistical Analysis with Missing Data*; Wiley and Sons: New York, NY, USA, 2002.
33. Demnati, A.; Rao, J.N.K. Linearization variance estimators for survey data. *Surv. Methodol.* **2004**, *30*, 17–26.
34. Prášková, Z.; Sen, P.K. Asymptotics in finite population sampling. In *Handbook of Statistics (Sample Surveys: Inference and Analysis)*; Pfeffermann, D., Rao, C.R., Eds.; Elsevier: Amsterdam, The Netherlands, 2009; Volume 29B, Chapter 40.
35. Frankovic, K.A.; Panagopoulos, C.; Shapiro, R.Y. Opinion and election polls. In *Handbook of Statistics (Sample Surveys: Design, Methods and Applications)*; Pfeffermann, D., Rao, C.R., Eds.; Elsevier: Amsterdam, The Netherlands, 2009; Volume 29A, Chapter 22.
36. Spagat, M. Suspicious Supervisors and Suspect Surveys. Available online: http://www.stats.org (accessed on 15 June 2016).
37. Jarmin, R.; Lane, J.; Marco, A.; Foster, I. Using the Classroom to Bring Big Data to Statistical Agencies. *Amstat News*, November 2014; pp. 12–13.
38. Li, L. Statistics Training: A Big Role in Big Data? *Amstat News*, May 2014, pp. 32–33.
39. Van der Laan, M. Statistics as a Science, Not an Art: The Way to Survive in Data Science. *Amstat News*, February 2015, pp. 29–30.
40. Priestley, J.L. Data Science: The Evolution or Extinction of Statistics? *Amstat News*, January 2016, pp. 26–27.
41. Varoufakis, Y. *Europe after the Minotaur: Greece and the Future of the World Economy*; Zed Books: London, UK, 2015.