Productivity Stagnation and Low Human Capital Investment in a Wealthy Economy: The Case of Italy

By Carlo Milana

During the last two decades, Italy has lagged behind in production and productivity growth with respect to other rich countries. This seems to be in contrast with its notable elements of wealth and the convergence predictions of modern growth theories. This work aims at identifying critical features of this puzzling case in a variant version of Solow’s growth accounting where “true” index numbers derived from Afriat’s approach allow for scale (dis)economies, (mis)allocation effects as well as technical (in)efficiency. The empirical results suggest that the relatively high technical and allocative inefficiencies in the Italian production system may be alleviated by eliminating the chronic vicious circle between underinvestment in human capital and depressingly low premiums in high-skilled labour rewards.

Italy’s productivity performance is strikingly different from that of other rich countries, which have themselves suffered a severe economic slowdown during the recent great recession. While some of these countries seem to recover the pre-crisis path of growth, Italy continues to show a worrisome tendency to fall behind in the process of economic growth. An important consequence of this behavior is of special importance. Productivity stagnation creates a vicious circle by reducing the real factor rewards and the country’s capacity to invest in further growth.

In order to help identifying key aspects of Italy’s productivity problem, this paper follows Afriat’s approach in a variant version of Solow’s growth accounting model. TFP growth is decomposed into technical change, scale effects, heterogeneity and misallocation, which are particularly suitable to investigate Italy’s productivity problems while imposing minimal assumptions on this analytical exercise.

1 J.E.L. classification codes: D24, E21, E22, E23, E24, E25, E26, J24, O47. This paper has been prepared for an invited presentation to the third World KLEMS conference on the theme “Growth and Stagnation in the World Economy” held by RIETI and Harvard University, Department of Economics in Tokyo, Japan on 19-20 May, 2014.

2 Birkbeck College, University of London. Email: c.milana@bbk.ac.uk. Personal email: carlo.milana@iol.it.
Previous empirical results have shown how the country has chronically suffered from structural problems and misguided policies, low investments in human capital and infrastructures coupled with tight fiscal measures. These factors are particularly difficult to adjust in the short and medium run to the need of structural changes. In an increasing globalization of the world economy, many rich countries endeavor to reinforce their comparative advantages in high profile and fast growing activities while primary stages of their supply chains are transferred abroad.

In this context, Italy is particularly disadvantaged with respect to other major developed economies. Indeed, the economy is endowed with a relatively high level of tangible and financial assets. Among the European major countries, Italy is among the richest in terms of real and financial capital although with low internal returns (Table 1 and Figure 1). In terms of total debt, the country features a low debt/GDP ratio (Figure 2). However, during the last two decades, the country has been scarcely able to offset the loss in the domestic production of mature industries with new activities. This fact can be attributed to the low presence of high-skilled labour (Figure 3), which is compensated less than the medium-skilled labour (Figure 4). The peculiar contrast between a relatively high wealth in tangible assets and low level of intangible assets, such as high-skilled human capital, is often regarded as a characteristic weakness of the Italian economy (Mas, Milana, and Serrano 2012; Milana, Nascia, and Zeli 2013).

Italy’s productivity stagnation and depressed investments in intangible capital can be analysed with reference to recent economic literature. Institutions and policies as well as the cost and benefits of acquiring high quality education can affect the acquisition of high skills (see, for example, Acemoglu and Zilibotti 2011 and Acemoglu and Dell 2010). Empirical studies support the importance attributed to human capital as a growth factor. For instance, Chatzimichael and Tzouvelekas (2014) find that improvements in human capital explain a significant portion of the increase in labour productivity in a large sample of countries.

Certain frictions regarding the remuneration of labour and capital within the firms explain the persistence of the low presence of high quality labour even in situations where their marginal products can be potentially high (see, for example, Frank, 1984 and, more recently, Jovanovic 2014). The recent literature has pointed the reduction effect on productivity level of the idiosyncratic policy distortions that divert financial resources from firms that are more productive in favour of less efficient government protected firms and
institutions (Foster et al. 2001, 2006, 2008, Hsieh and Klenow 2009, 2010; Fernald and Neiman 2011; Bartelsman, Haltiwanger, and Scarpetta 2013; Restuccia and Rogerson, 2008, 2013; Brandt, Tombe, and Zhu 2013, Midrigan and Xu 2014, Jovanovic 2014). However, this literature has focused mainly on misallocation of resources between firms within industries overlooking important effects of resource misallocation that take place also within production units and between industries.

An overall assessment of the growth reducing effect of misallocation and scale diseconomies can be carried out within the framework of TFP growth accounting using an appropriate methodology. Barro (1999) emphasized the complementarity of the information derived from growth accounting and behavioural (endogenous) growth models. As any growth accounting, however, the analysis proposed here does not offer behavioural descriptions of new long-run equilibrium paths (Barro and Sala-i-Martin 1992, 1995, p. 352). Future trends and factor adjustments induced by TFP changes are out of the scope of the present study.

The remainder of the paper proceeds as follows. Section 1 presents the methodology of TFP growth accounting based on a modified Solow’s model in the presence of scale and misallocation effects. Section 2 describes the data and their sources. Section 3 presents the empirical results obtained at industry and economy levels. The last section concludes.

### Table 1. Nominal tangible assets per capita and internal rates of return in the overall economy

<table>
<thead>
<tr>
<th></th>
<th>USA</th>
<th>Japan</th>
<th>U.K.</th>
<th>Germany</th>
<th>Italy</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Nominal tangible assets per capita (USA = 100)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1995</td>
<td>100.0</td>
<td>158.0</td>
<td>73.5</td>
<td>105.0</td>
<td>113.1</td>
</tr>
<tr>
<td>2000</td>
<td>100.0</td>
<td>155.0</td>
<td>71.0</td>
<td>109.0</td>
<td>120.0</td>
</tr>
<tr>
<td>2005</td>
<td>100.0</td>
<td>154.0</td>
<td>69.6</td>
<td>119.6</td>
<td>133.5</td>
</tr>
<tr>
<td>2010</td>
<td>100.0</td>
<td>150.0</td>
<td>67.0</td>
<td>117.0</td>
<td>129.0</td>
</tr>
<tr>
<td><strong>Gross internal rate of return (percent)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1995</td>
<td>6.3</td>
<td>4.2</td>
<td>5.7</td>
<td>4.6</td>
<td>3.6</td>
</tr>
<tr>
<td>2000</td>
<td>8.5</td>
<td>4.9</td>
<td>5.9</td>
<td>6.9</td>
<td>5.9</td>
</tr>
<tr>
<td>2005</td>
<td>11.8</td>
<td>6.0</td>
<td>6.7</td>
<td>7.7</td>
<td>5.3</td>
</tr>
<tr>
<td>2010</td>
<td>7.0</td>
<td>4.7</td>
<td>5.0</td>
<td>4.5</td>
<td>4.0</td>
</tr>
</tbody>
</table>

Source: Computations on EUKLEMS data (updated to 2010 using N.A. statistics)
Source: Bank of Italy (2008) and Cinquegrana (2010).

Source: See Figure 1.
Source: Computations on EUKLEMS data (updated using national statistics)
1. Methodology

The recent literature on growth accounting for developed and emerging economies alike has pointed out important effects of resource misallocation effects and scale diseconomies on TFP growth. The classical contributions in the field of taxation and monopoly regulation (Debreu 1951, 1954; Koopmans 1950; Herberg and Kemp 1971, Jones 1971) have indicated that distortions of various kinds in factor markets along with rationed resources can bring about non-trivial output losses (see Chandler and Miller, 2014 for an updated review). In order to account for short run disequilibrium, Morrison (1988, 1989, 1990) has developed a variant of Solow’s growth accounting using short-run econometrically estimated cost functions for manufacturing industries in the U.S., Canada, and Japan. Following this approach, Rossi and Toniolo (1999) have found that Italy’s pre-war productivity performance has been affected by scale diseconomies and lagging adjustments in quasi-fixed capital inputs.

Crafts and Mills (2005) have obtained similar results for the U.K. and Germany regarding both early and recent post-war periods. In the case of these two countries, the assumptions of marginal product pricing, constant returns to scale and instantaneous adjustments to long run cost equilibrium imposed on the traditional Solow’s accounting framework may lead to substantial biased conclusions regarding the interpretation of measured TFP as a technology indicator.

However, the above-mentioned measures are subject to the joint maintained hypotheses regarding the functional form and the stochastic component of the model (Afriat 1972, and Russell’s 1992). Another serious drawback may arise from the failure of duality consistency in the functional forms used in the primal and dual approaches to TFP measurement based on quantities and prices, respectively. This type of inconsistency may lead these two alternative approaches to different results3. This inconsistency is burdensome particularly in the inspection of the distribution of productivity gains among the real factor rewards.

3 Another discrepancy, which in fact is external to the primal-dual comparison per se, is typically obtained by using alternative prices implying quantities that are different from those actually used (see Hulten’s, 2009, p. 18, fn. 19 remarks).

1.1 Definitions

Let us define a set of price and quantity data for a production scheme in a number of observation points with $y_t$ and $p_t$, respectively the output quantity and producer price (net of unit margin of pure profits over marginal cost), and $x_t$ and $w_t$, respectively $N$-order vectors of input quantities and prices. These data can be discussed and analysed to assess whether and in what degree they can be rationalized in terms of dual production and minimum cost functions associated with an optimizing producer behaviour. More specifically, in order to achieve meaningful aggregating index numbers, the dataset is preliminary tested to be fully or partially consistent with the convexity of the quantity set $(x_t) \in \Omega^n$ and a non-decreasing and quasi-concave production function $f_{(t)}(x_t) = \max\{y_t : (x_t, y_t) \in \Omega^{n+1}\}$ with a possible additional homothetic character needed to define $TFP$ univocally. Under this condition, the input-output combinations could be rationalized by the inequality $y_t \leq f_{(t)}(x_t)$.

The dual function of minimum cost for which the output $y_t$ could be produced at input prices $w_t$ is

$$C_{(t)}(w_t, y_t) \equiv \min_x \{w \cdot x : f_{(t)}(x) \geq y_t\}$$

and therefore,

$$C_{(t)}(w_t, y_t) \leq C_{(t)}(w_t, f_{(t)}(x_t)) \leq C_{(t)^*}$$
where $C_t^* = w_t \cdot x_t$ is the observed total cost of production and, correspondingly,

$$y_t \leq f_{(t)}(x_t) \leq y(w_t, C_t^*)$$

where $y(w_t, C_t^*)$ is the return output function derived from inverting the minimum cost function (1) at the given level of $C_t^*$.

Cost efficiency and efficacy are measured by the extreme ratios:

$$e^{D}_t(w_t, x_t, y_t) = C_{(t)}(w_t, y_t) / C_t^*,$$

$$e^{E}_t(w_t, x_t, y_t) = y_t / y(w_t, C_t^*)$$

which are the products of intermediate ratios, respectively defining technical and allocative efficiencies on the side of output and input as

$$e^{D}_s(w_t, x_t) = C_{(t)}(w_t, f_{(t)}(x_t)),$$

$$e^{D}_m(w, x) = C_{(t)}(w, f_{(t)}(x)) / C_t^*$$

and defining technical and allocative efficacies on output and inputs as

$$e^{E}_s(y, x_t) = y_t / f_{(t)}(x_t),$$

$$e^{E}_m(w, x) = f_{(t)}(x_t) / y(w_t, C_t^*)$$

The indexes (4)-(6) are obviously the upper bounds of the cost efficiency.

From (5), we derive

$$e^{D}_m \cdot w_t \cdot x_t = \rho(w_t, x_t)$$

where $\rho(w_t, x_t) \leq C_{(t)}(w_t, f_{(t)}(x_t))$ with $e^{D}_m \leq e^{D}_m(w, x)$.

Using Afriat’s concept of critical cost-efficiency as the minimum cost efficiency across all observation points, that is $e^{E}_m = \min_{t} e^{E}_m$, we have for all $t$'s
(8) \[ e^*_M \cdot w_t \cdot x_t \leq \rho(w_t, x_t) \]

**Remark 1:** All data are fully consistent with rational (cost-minimizing) producer’s behaviour if the critical allocative efficiency index is equal to unity \( (e^*_M = 1) \). If \( e^*_M < 1 \), then the data are said to be consistent with partial allocative cost efficiency.\(^5\)

TFP measurement is meaningful, if and only if the producer’s technology is subject to the input-output separability and aggregability conditions, under which the production function becomes

\[
y_t = A_t F(\phi(x))
\]

(9) where \( A_t \equiv e_{t(\cdot)} \cdot a(t) \cdot a(t) \).

The function \( \phi(x) \) is a “true” aggregator of input quantities, and in particular it is quasi-concave and homogeneously linear (for any real number \( \lambda \), \( \phi(\lambda x) = \lambda \phi(x) \)). The associated cost inequality is

\[
e^D_{A_t} = \theta(w_t) \cdot \phi(x_t) \cdot \frac{1}{w_t \cdot x_t}, \quad 0 \leq e^D_{A_t} \leq 1
\]

(10) The index \( e^D_{A_t} \) can be partitioned into heterogeneity and misallocation components, that is \( e^D_{A_t} = e^H_{A_t} \cdot e^D_{A_t} \cdot e^H_{A_t} \cdot e^D_{A_t} \), and \( e^D_{A_t} \leq e^D_{A_t} \). The critical heterogeneity and allocation cost efficiency is then computed as \( e^*_A = \min_{t} e^D_{A_t}(w_t, x_t) \) to obtain

\[
e^*_A \cdot w_t \cdot x_t \leq \theta(w_t) \cdot \phi(x_t), \quad 0 \leq e^*_A \leq 1
\]

(11) with strict equality for at least one \( t \).

**Remark 2:** All data are fully consistent with rational (cost-minimizing) behaviour subject to an homothetic production function if the critical allocative efficiency and homotheticity index

\[^4\text{It is noteworthy that in equation (8) } t \text{ is not argument of } \rho(w_t, x_t).\]

\[^5\text{Some authors have proposed to achieve consistency with full allocative cost efficiency by deleting problematic data.}\]
is equal to unity ($e_A^* = 1$). If $e_A^* < 1$, then the data are said to be consistent with partial allocative cost-efficiency and homotheticity.

Based on (10), the observed cost can be partitioned into price and quantity components, that is $w_i \cdot x_i = W_t \cdot X_t$, where

$$W_t = \theta(w_i), \quad X_t = \phi(x_i) \cdot \frac{1}{e_A^*}$$

Re-writing (9) more explicitly as $y_i = A_t \cdot (X_t, e_A^*)^{\epsilon}$ with the associated dual cost function $C_t(w_t, y_t) = A_t \cdot \frac{1}{\epsilon} \cdot W_t \cdot y_t^{-\epsilon} \cdot e_A^{-1}$, in view of (10) and (12), the primal and dual measures of TFP level are the following

$$TFP^p \equiv \frac{y_t}{X_t} = \frac{A_t \cdot (e_A^*)^{\epsilon}}{X_t},$$

and

$$TFP^d \equiv \frac{W_t}{C_t \cdot y_t} = A_t \cdot \frac{1}{\epsilon} \cdot y_t^{-\epsilon} \cdot e_A^{-1}$$

**Remark 3:** Heterogeneity and input misallocation revealed by the inequality $e_A^* < 1$ reduce the measured TFP level.

**Definition 1:** Decomposition of the primary measure of TFP level and growth. The traditional primary measure of TFP level can be decomposed as follows:

$$TFP^p \equiv \frac{y_t}{X_t} = \frac{A_t \cdot \phi(x_i)^{\epsilon}}{\phi(x_i) / e_A^*} \quad \text{with} \quad A_t \equiv e_A^* \cdot a_{t} = \frac{y_t}{\phi(x_i)^{\epsilon}}$$

$$= \frac{y_t}{\phi(x_i)^{\epsilon} \cdot \phi(x_i)^{\epsilon} \cdot e_A^*}$$

and, the corresponding TFP growth index number between period 0 and period 1 is decomposed as follows
\( TFP_{1,0}^p = \frac{Y_{1,0}}{X_{1,0}} = TC_{1,0} \cdot SE_{1,0} \cdot OC_0 \cdot e_{A1} \)

where

\( Y_{1,0} = y_1 / y_0 \): output quantity index number;

\( X_{1,0} \): Input quantity index number (including inefficiency in input allocation);

\( EX_{1,0} = \phi(x_1) / \phi(x_0) = e_{A1} / e_{A0} \cdot X_{1,0} \): "True" index number of efficiently used input quantities;

\( TC_{1,0} \equiv A_1 = \frac{y_1}{y_0} \cdot (EX_{1,0})^{e_{A0}} \): “True” index number of technical change\(^6\),

\( \varepsilon_{1,0} \equiv \alpha_{1,0} e_{A0} w_{0,0} x_{0,0} + (1 - \alpha_{1,0}) e_{A1} w_{1,0} x_{1,0} / p_0 y_0 \) is the measured marginal elasticity of scale, and \( \alpha_{1,0} \) is the weight to be calibrated in the interval \( 0 \leq \alpha_{1,0} \leq 1 \).

\( SE_{1,0} = (EX_{1,0})^{e_{A0}^{-1}} \): “True” index number of scale effect taking the numerical values;

\( OC_0 \equiv \frac{1}{e_{A0}} \): Potential gain in TFP level from an organizational change offsetting the input-cost inefficiency in period 0.

**Definition 2: Decomposition of the dual cost-based measure of TFP level and growth.**

The traditional dual cost-based measure of TFP level can be decomposed as follows:

\( TFP_{1,0} \equiv \frac{TFP}{TFP_0} = \frac{W_1 / W_0}{C^*_1 / C_0} \cdot \frac{y_1}{y_0} = \left( \frac{A_1}{A_0} \right)^{1/e_{A0}} \cdot \left( \frac{y_1}{y_0} \right)^{1/e_{A0}} \cdot \frac{1}{e_{A0}} \cdot e_{A1} \)

\(^{6}\) The index of technical change, in turn, could be accounted for with changes in the benchmark technological frontier, scale economies (SE), and technical efficiency (TE) if further information regarding at least one of these two last components is brought into the picture, as for example in Färe, Grosskopf, Norris, and Zhang (1994, 1997), and Ray and Desli (1997).
1.2 Input aggregation and heterogeneity/misallocation effects

The initial proposition developed by Samuelson in his revealed preference theory was later completed with the following theorem:

**Afriat’s (1967) Theorem:** The three conditions of cyclical, multiplier, and level consistency on the cross-structure of an expenditure configuration are all equivalent, and are implied by the condition of utility for the configuration.

Afriat’s *cyclical condition*, later called Generalized Axiom of Revealed Preference (GARP) by Varian (1982, 1984), is a generalized form of Houthakker’s strong axiom of revealed preference and involves only observed market data on prices and quantities. This was a major advance for empirical analysis being equivalent to the other two conditions, the multiplier and level consistency, which involve non-observable utility levels and marginal utility of income. The formulation of the cyclical conditions using Afriat’s inequalities would enable the elimination of all the unknowns in the construction of index numbers.

The cyclical (or GARP) condition implies that the available demand data can be rationalized with a cost minimizing behaviour, and a set of piece-wise linear minimum cost functions \( \rho(w, x) \), if there is no binary comparison \((t, s)\) exhibiting a Paasche index increase, \( w_i x_t / w_j x_t > 1 \) in contrast with a Laspeyres index decrease, \( w_i x_t / w_j x_t < 1 \).

If the cyclical (or GARP) test is violated, however, it is still possible to find a rationalizing behaviour at a lower degree. Afriat (1972) established a way to weaken the test in order to allow for some inefficiency by introducing the index, in our notation \( e_{Mt} \) taking values in the interval \( 0 \leq e_{Mt} \leq 1 \) (called Afriat efficiency index by Varian, 1990, 1992) with \((1-e_{Mt})\) being the relative inefficiency index. Among all observations \( t \)'s, an index of *critical cost-efficiency* \( e_{Mt}^* \) can be found as the lowest among the maximum values \( e_{Mt} \) ensuring that the cyclical (or GARP) test be not violated. This implies the existence of a set of minimum
cost functions $\rho(w, x)$ that are consistent with a partially rationalizing producer’s behaviour as shown in equation (8).\(^7\)

In another path-breaking work, Afriat (1972, p. 49, 1981, p. 155) examined the constructability of consistent (transitive chained) index numbers between several periods simultaneously by redefining his theorem under more stringent conditions, which Varian (1983, p. 103) later called Homothetic Axiom of Revealed Preference (HARP), as follows:

**Afriat’s (1972, p.49, 1981, p.155) homogeneity theorem:** *If and only if, for every sequence $i,k,...,l,m$, the data set satisfies the following inequality between chained Laspeyres indexes and the direct Paasche index*

\begin{equation}
X_{ij}^L X_{jk}^L ... X_{lm}^L X_{mi}^L \geq 1
\end{equation}

*or, equivalently, since $X_{mi}^L = 1/ X_{im}^P$,*

\begin{equation}
X_{ij}^L X_{jk}^L ... X_{lm}^L \geq X_{im}^P
\end{equation}

*where $X_{st}^L$ is a Laspeyres quantity index and $X_{im}^P$ is a Paasche quantity index, then there is at least one homothetic production function rationalizing the data enabling the construction of the upper and lower limits of homogeneous “true” quantity and price index numbers.*

Testing HARP conditions have become a standard procedure to establish if a particular demand data set is fully consistent with separability and aggregability conditions without imposing any joint hypothesis concerning a specific functional form of the rationalizing function.\(^8\)

---

\(^7\) The cost-efficiency index can be decomposed into a determinist component $e_{st}$ due to misallocation effects and a stochastic component $\exp^s$ where $u_t$ has the distribution $N(0, \sigma)$. Hence, in this case $e_{st} = e_{st} \cdot \exp^s$. The introduction of the stochastic term has in fact given way to the literature on the stochastic frontier analysis imposing, however, the hypothesis of a specific function form to be econometrically estimated.

\(^8\) It may be useful to take note, however, that the fact that the data set satisfies HARP does not exclude that the same data could be also consistent with or generated by some non-conical (non-homothetic) utility (or production) functions if a rational behaviour exists at all.
As shown independently by Shephard (1953, p. 43) and Afriat (1953-1956, 1972, p. 23), under input-output homothetic separability, a minimum cost function $\rho(w_t, x_t)$ can be factorized into a product of dual conjugate functions of price and quantity levels\(^9\), that is

\begin{equation}
\rho(w_t, x_t) = \theta(w_t) \cdot \phi(x_t)
\end{equation}

If Afriat’s homogeneity (or HARP) test is not passed, then the computations are corrected for the allocative cost-inefficiency index $e_{At}$. This is equivalent to a further weakening of the cost-efficiency parameter to be applied to the observed expenditure data, in order to obtain equations (10) and (11).

With the provision of a suitable correction for allocative inefficiency, if needed, the data always pass the HARP test. Then, along with efficiency index, the tight upper and lower limits of the “true” indexes of price and quantity levels can be found by tracking the minimized chained Laspeyres and maximized Paasche index numbers, respectively, using a computation algorithm. Appendix A describes the computation algorithm proposed by Afriat (1981) himself, which has been used in the present work. The obtained cost-inefficiency index is further decomposable using econometric techniques similar to those of the stochastic frontier approach.

1.2. Technical change and scale (dis)economies

A vast economic literature offers a wide range of anecdotes and reasons why scale (dis)economies may appear in production activities although many authors prefer to advocate the constant returns to scale hypothesis (which is indeed realistic at least in the long-run equilibrium). Disentangling empirically the scale effects on output from those of technical change using a representative production function has been a classical problem for decades. Solow (1957) noted that he was “... using the phrase ‘technical change’ as a shorthand expression for any kind of shift in the production function” (italics in the original). The famous controversy between Stigler (1961) and Solow (1961) regarding whether a homogeneous production function with Hicks-neutral technological change offers a more

---

\(^9\) The concepts of aggregate level indexes have been theoretically defined by Keynes (1909, 1930), Shephard (1953), Afriat (1953-1956,1972), Samuelson and Swamy (1974, p. 570).
realistic representation than a homothetic function with no technical progress but non-
constant returns to scale was part of the early debate. A definite clarification has been
provided by Sato (1980), and Sato and Calem (1983) (see also Sato and Ramachadran, 2014),
who have indicated that both interpretations were valid as the Hicks-neutral technical
progress is indistinguishable from homothetic scale effects making this identification problem
impossible to solve within this analytical framework. Other authors have extended this
impossibility theorem to the case of biased technical change when also the elasticities of
substitution are to be simultaneously measured (Diamond, McFadden, and Rodriguez, 1978,
and Greene, 1983)\textsuperscript{10}.

It turns out however that productivity measurement can gain additional information
from the simultaneous inspection of the dual TFP accounting of costs and the formation of
output prices. The inverse of the ratio of the estimated marginal cost over total costs is in fact
the local output elasticity with respect to aggregate inputs. The approach adopted by Caves,
Christensen, and Diewert, 1982 and Chan and Mountain (1983), recently extended to the
multiple-output case by Diewert and Fox (2010, 2013), and Boles de Boer and Evans (1996)
includes external information on the marginal cost pricing in imperfectly competitive
markets\textsuperscript{11}. Hsieh and Klenow (2009) use a similar variable in the case of China, by
interpreting the current market regimes through the prism of monopolist competition.

We assume imperfectly competitive and inefficient markets, where producers are
price takers in input markets and, given the size of the economy, manage to control the output
price formation only with a very limited power. The output quantity elasticity in time \( t \) with
respect the efficiently used inputs can be obtained as follows

\begin{equation}
\epsilon_t \equiv \frac{\partial y_t}{\partial \phi(x_t)} \cdot \frac{\phi(x_t)}{y_t} = \frac{W_t \cdot X_t}{p_t \cdot y_t}, \quad \left. \frac{AC_t}{MC_t} \right|
\end{equation}

\textsuperscript{10} Closely related is the impossibility of carrying out tests on output and input separability by means of
econometric estimations independently from the assumption imposed \textit{a priori} on the particular functional form
of the model used (Blackorby, Primont, and Russell, 1977, and Denny and Fuss, 1977).

\textsuperscript{11} Alternatively, Diewert, Nakajiama, and Nakamura (2011) regressed the log-values of the Törnqvist output
index number on the log-values of the Törnqvist input index number using establishment-level data in order to
identify the rate of change of TC and the scale elasticity over any pair of time periods.
where $W_t$ is the index of input price level, the producer price $p_t$ is assumed to be set equal to marginal cost $MC_t = p_t^m(1 - m_t)$, the market price $p_t^m$ net of pure profit margin $m_t$, and $AC_t = W_t \cdot X_t \cdot e_t / y_t$ is the average cost of efficiently used inputs. With the average and marginal costs increasing with output level due to scale diseconomies, profit optimization would tend to fix the output level at a point where $MC_t \geq AC_t$, leading to positive profits. When $e_t \neq 1$, the TFP component due to the cost-revenue gap is consistent with a scale effect rather than a Hicks-neutral technical progress given that the real factor rewards do not fully internalize the output revenue.

The notion of economies or diseconomies of scale is not clearly stated. In his review article on the subject, Gold (1981, p. 5) recalled that “analyses have repeatedly called attention to the fuzziness of the basic concept of scale, and to uncertainties about the sources of expected benefits, as well as the relatively modest gains apparently derived from additional increases in scale”. Other authors have pointed to a number of different sources of scale of internal and external economies associated with the scale of production. Meade (1952) considered two relevant types of economies or diseconomies external to the industry. The first type comes from “unpaid factors of production”, whereas the second comes from the “creation of atmosphere”. The essential difference between these two types is that in the first case there are constant returns to scale for the society as a whole, but not for the individual industry, whereas in the second case there may be constant returns to scale for the industry but not for the society as a whole. Stigler (1961) and Solow (1961) agreed on Marshallian externalities with economies of scale external to the firm but internal to the industry.

Regarding scale diseconomies of “external” origin, Barro and Xala-i-Martin (1992, 1995, pp. 158-161) considered growth models where public services are explicit inputs of production. Highways, water systems, police and fire services, courts and many other government services are subject to congestion leading to higher costs as production grows. Barro (1999, p. 126) mentions negative spillover effects, such as those from traffic congestion and environmental damage. Positive scale economies have been traditionally associated with growth driven by innovation. More recently, Raimondo, Rodríguez-Clare, and Sabario-Rodríguez (2012) have however showed that domestic frictions, rather than

---

12 Marshallian externalities are discussed by Prendergast (1993). A discussion of the earlier neoclassical debates about returns to scale, costs, and long-run supply is provided by Aslanbeigui and Naples (1997).
openness to trade and multinational production account for exceptions to the counterfactual prediction.

Having measured the scale elasticity with respect to the aggregate inputs (net and gross of allocative inefficiency), we are ready to construct index numbers of technical change ($TC_{1,0}$) and scale (dis)economies ($SE_{1,0}$) as defined within the TFP growth accounting equation (16). The scale effect component may take the numerical values shown in Table 1 where different cases of returns to scale and changes in input quantities are considered.

### Table 2. Scale (dis)economy effect on TFP relative level

<table>
<thead>
<tr>
<th>Scale effect</th>
<th>Index of inputs</th>
<th>Degree of return to scale</th>
</tr>
</thead>
<tbody>
<tr>
<td>$SE_{1,0} &gt; 1$</td>
<td>$EX_{1,0} &gt; 1$</td>
<td>increasing returns to scale</td>
</tr>
<tr>
<td></td>
<td>$EX_{1,0} &lt; 1$</td>
<td>decreasing returns to scale</td>
</tr>
<tr>
<td>$SE_{1,0} = 1$</td>
<td>Any (positive)</td>
<td>$EX_{1,0}$ and constant returns to scale</td>
</tr>
<tr>
<td></td>
<td>$EX_{1,0}$</td>
<td>any returns to scale</td>
</tr>
<tr>
<td>$SE_{1,0} &lt; 1$</td>
<td>$EX_{1,0} &gt; 1$</td>
<td>increasing returns to scale</td>
</tr>
<tr>
<td></td>
<td>$EX_{1,0} &lt; 1$</td>
<td>decreasing returns to scale</td>
</tr>
</tbody>
</table>

1.3. Aggregate productivity growth

Misallocation of resources occurring at micro-level within industries and firms reverberates throughout the economy in a way that can be properly tracked in an input-output framework (Jones 2011). This is implied by the aggregation of industry-level productivity growth using the well-known Domar’s (1961) weights.

Afriat’s procedure yields the usual measured decomposition of the nominal value of total final output of the economy

\[ \mathbf{p} \cdot \mathbf{q} = P(\mathbf{p}) \cdot Q(\mathbf{q}) \cdot \frac{1}{e^\theta} \]
where \( \mathbf{p} \) and \( \mathbf{q} \) are vectors of final-output prices and quantities of order \( n \) (the number of industries).

The inefficiency component \( e^b \) can be interpreted as the misallocation effect *between* industries. Hence, taking the vectors of log-change rates of industry-level TFP indexes and their components defined by (16), the log-change of the economy-wide TFP index is obtained and decomposed as follows

\[
\Delta \ln TFP_{1,0}^E = (\Delta \ln T + \Delta \ln OC + \Delta \ln SE_{1,0} + \ln e^*_t) \cdot \bar{d}_{1,0}
\]

(23)

where \( \bar{d}_{1,0} \) is the vector of average Domar’s weights \( d^i_j = \frac{p^i_j y^i_j}{\sum_j p^i_j q^i_j} \) for \( t = 0,1 \).

### 1.4 International comparisons

The international comparisons of price and volume levels of inputs and outputs are made using the metrics of values denominated in one single currency. The output prices in country \( r \) relative to the US prices prevailing in the base year, both expressed in US dollars, are given by

\[
\hat{B}_r = \frac{p^r y^r}{\hat{E} y}
\]

(24)

where \( \hat{PPP}_r \) is the vector of sectoral purchasing power parities in the base year for outputs, defined as ratios of sectoral output prices in country \( r \) expressed in domestic currency to the corresponding prices in the US expressed in dollars, \( E^{c,15} \) is the nominal exchange rate that converts values expressed in dollars into values in country’s \( r \)’s currency.

The volume levels of outputs \( y \) are converted in monetary values at constant US prices \( y^s \) using the same purchasing power parities at the base year:

\[
y^s = \hat{PPP}_B^{-1} \cdot y
\]

(25)

Multiplying (24) and (25) consistently yields \( p^s \cdot y^s = \frac{1}{E^{c,15}} \cdot p \cdot y \).
An analogous procedure can be adopted for the construction of sectoral price and volume levels of factor inputs in a country $r$ relative to those in the US prevailing in the base year, both expressed at US dollars.

2. The data

The data used in this study have been constructed using the EUKLEMS database for the years 1990-2005 (Van Ark, O’Mahony, and Timmer 2008; O’Mahony and Timmer 2009). These data have been extended to the years 2006-2010 using the input-output tables of the WIOD database and the national accounts statistics of U.S.A., Canada, Japan, U.K., France, and Italy. The sectoral purchasing power parities have been derived from the PPP statistics available on OECD.Stat using the traditional procedure of “peeling off” deductible indirect taxes and, for the output PPP, also transport and commercial margins to convert market consumer prices into \textit{ex fabrica} producer prices.

In particular, the following data have been used for 36 industries and other 15 sectors in which the total economy of each country has been subdivided:

1. \textit{Gross Outputs:} Volumes have been valued in US dollars at US prices prevailing in 2005, whereas the \textit{ex fabrica} output prices (net of transport and commercial margins) were expressed in US dollars in relative levels with respect to the corresponding US prices prevailing in 2005.

2. \textit{Capital inputs:} Capital volume inputs have been expressed into volumes of capital input services valued in US dollars at US (ex ante) rental prices prevailing in 2005, whereas the domestic (ex ante) rental price indexes are denominated in US dollars relative to the corresponding US rental prices prevailing in 2005. The capital inputs are distinguished in ICT and non-ICT categories.

3. \textit{Labor inputs:} Labour quantity inputs are measured in hour worked whereas the corresponding domestic factor prices are the resulting domestic labor compensation per hour denominated in US dollars. Labor inputs are distinguished in hours worked by high-skilled, medium-skilled, and low-skilled workers.

4. \textit{Intermediate inputs:} Volumes have been valued in US dollars at USA prices prevailing in 1995, whereas the domestic purchaser input prices (net of transport and commercial margins) are denominated in US dollars in relative levels with respect to
the corresponding US prices prevailing in 2005. The intermediate inputs are subdivide into three subaggregates of inputs: energy, materials, and services.

We adopt the most general assumption that all outputs and factor inputs are industry-specific. Therefore, no aggregation across industries is warranted for the same kind of input (for example, a plough used in Agriculture or a physician employed in the Health sector cannot be technically moved into fully productive activities of other industries without costly reconversion).

3. **Empirical results**

The index-number methodology built on Afriat’s approach and applied to a generalized Solow’s TFP growth accounting enables us to unveil interesting features of the Italian productivity malaise. They can be summarized in the following points:

1. During the years 1991-2010, the overall economy seemed to stay behind in productivity levels when other major European countries, namely U.K. and Germany, continued their process of convergence with the U.S. The U.S. economy remains the most dynamic in terms of technical change even though diseconomies of scale dampened somewhat the TFP growth. All the examined European countries remain below the US TFP level. Italy, however, is the only country suffering a stagnating TFP and a practically absent technical change at the aggregate level (Figures 5 and 6).

2. The most dynamic industries with an average growth rate in TFP and technical change higher than an average of 1 per cent per year are the ICT sector and Agriculture. But also some service industries such as Sale, maintenance, and repair and Public administration have achieved relatively high productivity performance under tight budget constraints. Important sectors such as Education have registered no change in both TFP and technical change. On the other side of the industry spectrum, Coke and refined petroleum products, Construction, and other services, such as Community services, Personal services, and Finance, have significantly reduced TFP and scored at the same time a negative technical change (Table 3 and Figure 7).

3. A wide variation of Italian industries can be noted in the productivity performance with respect to the U.S. counterparts as summarized with a variance higher than 30 percent and a coefficient of variation around 78 per cent (Figure 8).
4. In Italy, the scale and misallocation effects remained stable over the period and, therefore, the organizational efficiency has not significantly improved in the overall economy (Table 4 and Figure 9).

5. Scale effects on productivity are relatively low compared to the other examined countries including the U.S. This can be due to the depressed profile of economic growth that has not allowed possible scale (dis)economies to emerge (Figure 9).

6. The well-known undercapitalization in human capital of the Italian industries has contributed to heavier misallocation effects on productivity than in other rich countries. As already noted, the cost efficiency index lower than the unit level reveals heterogeneity in comparisons of the internal structure of production. Indeed, heterogeneity could arise with non-neutral technical change and scale effects, but in the case of Italy both these changes has been mild in the more recent examined period. Misallocation remains the main probable factor of the noted low cost efficiency in many Italian industries. The country’s allocative efficiency was much worse than other major EU members with more than 7 percent of output loss against 4 percent in Germany, 4.5 percent in U.K., and 5 percent in France (Figure 10). Among the industries that have performed with the least misallocation effects, the ICT sectors, Electrical equipment, Food and beverages, Basic metals, and Wood products scored output losses below 2.5 per cent. The industries that exhibited the worst performance in allocative efficiency are Electricity, gas, and water with an output loss about 18 per cent, Sale, maintenance, and repair about 15 percent, Non-market services (14 per cent), Retail Trade, Renting, and Finance between 10 and 15 percent (Figure 11).

4. Conclusion
The puzzling case of the stagnating Italian economy characterized by a rich endowment of tangible and financial capital and a post-war history of fast productivity growth has been analyzed with an appropriate growth accounting method. The empirical results suggest that a vicious circle of low accumulation of high-quality factors of production and low relative rewards to skilled labor had depressing effects on growth. The whole economy was entrapped in a state of misallocation of resources, a lack of high profile human resources, and no apparent technical change.

The empirical results encourage the expectations that the relatively high level of tangible assets that are still present in many sectors could be more efficiently be associated
with an increased skilled labor and new entrepreneurial activities in order to promote innovation through inventiveness, ideas, and knowledge. A prospective growth-oriented strategy might profitably gain momentum in the presence of an augmented human capital, which in turn might bring the economy to a more rational use of the country’s rich resources.
Table 3. TFP and TC index numbers of Italian industries (USA 1995 = 100)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Agriculture, hunting, forestry and fish</td>
<td>41.8</td>
<td>47.2</td>
<td>52.5</td>
<td>51.9</td>
<td>51.8</td>
</tr>
<tr>
<td>Mining and quarrying</td>
<td>66.3</td>
<td>73.6</td>
<td>74.3</td>
<td>70.2</td>
<td>68.7</td>
</tr>
<tr>
<td>Food products, beverages and tobacco</td>
<td>77.0</td>
<td>79.5</td>
<td>79.5</td>
<td>76.4</td>
<td>76.7</td>
</tr>
<tr>
<td>Textiles, textile products, leather an</td>
<td>102.3</td>
<td>109.6</td>
<td>109.1</td>
<td>104.2</td>
<td>104.7</td>
</tr>
<tr>
<td>Wood and products of wood and cork</td>
<td>68.3</td>
<td>71.6</td>
<td>76.7</td>
<td>76.1</td>
<td>76.3</td>
</tr>
<tr>
<td>Pulp, paper, paper products, printing</td>
<td>78.1</td>
<td>81.7</td>
<td>81.5</td>
<td>79.5</td>
<td>79.4</td>
</tr>
<tr>
<td>Coke, refined petroleum products and n</td>
<td>54.0</td>
<td>57.0</td>
<td>49.6</td>
<td>43.9</td>
<td>45.9</td>
</tr>
<tr>
<td>Chemicals and chemical products</td>
<td>66.4</td>
<td>70.4</td>
<td>70.7</td>
<td>71.4</td>
<td>71.2</td>
</tr>
<tr>
<td>Rubber and plastics products</td>
<td>86.5</td>
<td>90.2</td>
<td>89.9</td>
<td>89.4</td>
<td>89.5</td>
</tr>
<tr>
<td>Other non-metallic mineral products</td>
<td>78.5</td>
<td>77.7</td>
<td>76.2</td>
<td>73.4</td>
<td>79.2</td>
</tr>
<tr>
<td>Basic metals and fabricated metal prod</td>
<td>77.8</td>
<td>84.5</td>
<td>84.1</td>
<td>83.9</td>
<td>83.8</td>
</tr>
<tr>
<td>Machinery, n.e.c.</td>
<td>90.6</td>
<td>95.1</td>
<td>91.3</td>
<td>91.9</td>
<td>91.7</td>
</tr>
<tr>
<td>Electrical and optical equipment</td>
<td>103.6</td>
<td>108.5</td>
<td>108.2</td>
<td>106.7</td>
<td>106.6</td>
</tr>
<tr>
<td>Transport equipment</td>
<td>87.9</td>
<td>87.8</td>
<td>90.3</td>
<td>86.8</td>
<td>87.3</td>
</tr>
<tr>
<td>Manufacturing n.e.c.; recycling</td>
<td>83.6</td>
<td>88.5</td>
<td>88.9</td>
<td>88.7</td>
<td>88.1</td>
</tr>
<tr>
<td>Electricity, gas and water supply</td>
<td>94.7</td>
<td>94.3</td>
<td>91.2</td>
<td>95.2</td>
<td>94.8</td>
</tr>
<tr>
<td>Construction</td>
<td>14.3</td>
<td>19.0</td>
<td>19.3</td>
<td>18.7</td>
<td>18.9</td>
</tr>
<tr>
<td>Wholesale trade, and commission trade</td>
<td>89.3</td>
<td>102.3</td>
<td>100.7</td>
<td>98.3</td>
<td>98.0</td>
</tr>
<tr>
<td>Retail trade, except of motor v.</td>
<td>51.2</td>
<td>66.0</td>
<td>64.7</td>
<td>64.0</td>
<td>63.8</td>
</tr>
<tr>
<td>Pulp, paper, paper products, printing</td>
<td>80.8</td>
<td>87.9</td>
<td>85.3</td>
<td>85.5</td>
<td>85.6</td>
</tr>
<tr>
<td>Post and telecommunications</td>
<td>52.0</td>
<td>51.8</td>
<td>60.9</td>
<td>75.5</td>
<td>75.0</td>
</tr>
<tr>
<td>Financial intermediation</td>
<td>96.1</td>
<td>96.4</td>
<td>93.6</td>
<td>93.5</td>
<td>93.5</td>
</tr>
<tr>
<td>Real estate activities</td>
<td>80.5</td>
<td>80.8</td>
<td>78.9</td>
<td>82.1</td>
<td>82.4</td>
</tr>
<tr>
<td>Renting of meq and other business</td>
<td>57.1</td>
<td>56.5</td>
<td>56.9</td>
<td>57.7</td>
<td>57.6</td>
</tr>
<tr>
<td>Hotels and restaurants</td>
<td>85.2</td>
<td>89.1</td>
<td>85.1</td>
<td>84.7</td>
<td>84.5</td>
</tr>
<tr>
<td>Public admin and defence</td>
<td>56.4</td>
<td>56.4</td>
<td>58.2</td>
<td>61.7</td>
<td>61.8</td>
</tr>
<tr>
<td>Construction</td>
<td>39.0</td>
<td>39.2</td>
<td>39.4</td>
<td>39.4</td>
<td>39.4</td>
</tr>
<tr>
<td>Total Payments</td>
<td>72.1</td>
<td>74.5</td>
<td>74.7</td>
<td>73.4</td>
<td>73.5</td>
</tr>
</tbody>
</table>

TOTAL ECONOMY: 72.1 74.5 74.7 73.4 73.5

Table 4. Scale effect and cost efficiency indexes in the Italian industries

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Agriculture, hunting, forestry and fish</td>
<td>1.000</td>
<td>0.975</td>
<td>0.972</td>
<td>0.969</td>
<td>0.970</td>
</tr>
<tr>
<td>Mining and quarrying</td>
<td>1.000</td>
<td>0.977</td>
<td>0.999</td>
<td>0.999</td>
<td>0.999</td>
</tr>
<tr>
<td>Food products, beverages and tobacco</td>
<td>1.000</td>
<td>0.971</td>
<td>0.951</td>
<td>0.931</td>
<td>0.950</td>
</tr>
<tr>
<td>Textiles, textile products, leather an</td>
<td>1.000</td>
<td>0.999</td>
<td>0.999</td>
<td>0.999</td>
<td>0.999</td>
</tr>
<tr>
<td>Wood and products of wood and cork</td>
<td>1.000</td>
<td>1.001</td>
<td>1.006</td>
<td>1.002</td>
<td>1.004</td>
</tr>
<tr>
<td>Pulp, paper, paper products, printing</td>
<td>1.000</td>
<td>0.999</td>
<td>0.995</td>
<td>0.996</td>
<td>0.996</td>
</tr>
<tr>
<td>Coke, refined petroleum products and n</td>
<td>1.000</td>
<td>0.999</td>
<td>0.997</td>
<td>0.995</td>
<td>0.998</td>
</tr>
<tr>
<td>Chemicals and chemical products</td>
<td>1.000</td>
<td>0.999</td>
<td>0.997</td>
<td>0.995</td>
<td>0.998</td>
</tr>
<tr>
<td>Rubber and plastics products</td>
<td>1.000</td>
<td>0.999</td>
<td>0.997</td>
<td>0.995</td>
<td>0.998</td>
</tr>
<tr>
<td>Other non-metallic mineral products</td>
<td>1.000</td>
<td>0.999</td>
<td>0.997</td>
<td>0.997</td>
<td>0.998</td>
</tr>
<tr>
<td>Basic metals and fabricated metal prod</td>
<td>1.000</td>
<td>0.999</td>
<td>0.997</td>
<td>0.997</td>
<td>0.998</td>
</tr>
<tr>
<td>Machinery, n.e.c.</td>
<td>1.000</td>
<td>0.999</td>
<td>0.997</td>
<td>0.997</td>
<td>0.998</td>
</tr>
<tr>
<td>Electrical and optical equipment</td>
<td>1.000</td>
<td>0.999</td>
<td>0.997</td>
<td>0.997</td>
<td>0.998</td>
</tr>
<tr>
<td>Transport equipment</td>
<td>1.000</td>
<td>1.007</td>
<td>1.016</td>
<td>1.014</td>
<td>1.013</td>
</tr>
<tr>
<td>Electricity, gas and water supply</td>
<td>1.000</td>
<td>0.977</td>
<td>0.991</td>
<td>0.987</td>
<td>0.985</td>
</tr>
<tr>
<td>Construction</td>
<td>1.000</td>
<td>0.977</td>
<td>0.991</td>
<td>0.991</td>
<td>0.991</td>
</tr>
<tr>
<td>Total Payments</td>
<td>1.000</td>
<td>0.996</td>
<td>0.976</td>
<td>0.969</td>
<td>0.971</td>
</tr>
</tbody>
</table>

TOTAL ECONOMY: 1.000 0.996 0.976 0.969 0.971
Figure 9. TFP growth against scale diseconomy effect on output in Italy, 1991-2010 (% p.a.)

Figure 10. TFP growth against output loss due to misallocation of resources within industries in Italy, 1991-2010 (% p.a.)
Figure 11. Average growth rates of TFP and TC and relative output loss due to misallocation, 1991-2010 (% p.a.)
APPENDIX A

Procedures of aggregation and cost efficiency measurement

1. Input Aggregation

In order to define the aggregate measure of input quantities, we refer to the recent discussion in this journal by van Veelen and van der Weide (2008). They noted that the traditional economic approach to index numbers imposes arbitrarily not only a chosen formula, but also the assumption that the data are originated by optimizing (fully efficient) choices. However, one of the problems with the traditional index number approach is that the imposed index-number formula is not neutral with respect to the cardinal (and even ordinal) ranking of the compared situations. The economic literature has debated this problem for a long time, especially in the international comparisons of incomes and productivity. Feenstra, Ma, Neary, and Rao (2013) have obtained significantly different measures of real per capita GDP with different index number formulas with remarkable differences in 2005 than that reported by the World Bank (2008) using 2005 ICP data.

Moreover, as already noted by Samuelson and Swamy (1974), certain axioms such as those regarding consistency in aggregation, transitivity, and linear homogeneity (together with a few other properties considered initially by Irving Fisher 1922) for the obtained quantity and price indexes are the minimum requirements for the indexes to provide meaningful economic measures. Unfortunately, the set of index number formulas satisfying these minimum requirements is generally empty as noted already by Frisch (1936) including the approximating formulas in vogue today. A similar impossibility theorem arose in the discussion of the axiomatic approach to multilateral index numbers by Van Veelen (2002).

For this reason, van Veelen and van der Weide (2008) carried out their discussion of the economic approach by following Afriat (1967, 1981), whose nonparametric method is not based on the choice of index-number formulas but uses a procedure based on a theorem derived on a reformulation of the revealed preference theory (Afriat 1964). This theorem stated that a piece-wise linear utility or production function exists which rationalizes the data if and only if a “cyclical consistency” of inequality conditions (corresponding to inequalities regarding direct Laspeyres and Paasche indexes) are satisfied. This path-breaking improvement makes it possible to recover implicitly a family of linear-piece wise utility or production functions rationalizing a finite dataset using appropriate price and quantity index numbers not based on fixed formulas.

Afriat’s vision, not noted for a long time in the economic literature with few exceptions including Hanoch and Rothschild (1972), Dievert (1973), Diwvert and Parkan (1983,1985), became eventually more widely known when his conditions on “cyclical consistency” were re-proposed under the name of General Axiom of Revealed Preference (GARP) by Varian (1982, 1983, 1984). If these

---

13 Hill (2006) has found a large spread in numerical values of superlative index numbers, with the largest and smallest ones differing by more than 100 per cent using a standard US national data set and by about 300 per cent in a cross-section comparison of countries using an OECD data set. The controversial debate on how to properly compute and interpret the aggregate price and PPP indexes for China and other emerging countries is well represented in the discussion between Ravallion (2013a, 2013b) and Inklaar (2013).

14 On the literature regarding the existence of a utility function and its distinction from consumer preferences, see for example Metha (1998). Many of those results are isomorphic in the field of producer’s choices (Varian, 1984).

15 Fostel, Scarf, and Todds (2004) give a somewhat more transparent reformulation of Afriat’s theorem. Empirical application of Afriat’s approach include Dowrick and Quiggin (1994, 1997), and Bar-Shira,
conditions are satisfied, then the data on \( n \) input quantities \( x \) purchased at prices \( w \) for a number of \( m \) periods are consistent with a rational economic behaviour. Therefore, a class of alternative minimum cost functions can be postulated. Each of these functions is dual to a respective utility (or production) function that rationalizes the data.

In another path-breaking work, Afriat (1972, p.49, 1981, p.155) examined the constructability of consistent (transitive chained) index numbers between several periods simultaneously by redefining his theorem under more stringent conditions that Varian (1983, p. 103) later called Homothetic Axiom of Revealed Preference (HARP). If and only if, for every sequence \( i,k,...,l,m \) the data set satisfies

\[ \text{Afriat (1981, p.155) applied this theorem to the constructability of consistent (transitive chained) index numbers between several periods simultaneously. Testing HARP conditions have become a standard procedure to establish if a particular data set is fully consistent with aggregations without imposing any further joint hypothesis concerning a specific functional form of the rationalizing utility (or production) function.}^{16} \]

As already shown independently by Shephard (1953, p. 43) and Afriat (1953-1956, 1972, p. 23), under input-output homothetic separability, a minimum cost function \( \rho(w_t, x_t) = C_{(i)}(w_t, y_i(x_t)) \) can be factorized into a product of dual conjugate functions of price and quantity levels\(^{17} \), that is

\[
(A1) \quad \rho(w_t, x_t) = \theta(w_t) \cdot \phi(x_t)
\]

for every observed period \( t \) or \( C_{(i)}(w_t, y_i) = \theta(w) \cdot F_{(i)}^{-1}(y) \).

The HARP condition imply the existence of the optimized chained Laspeyres and Paasche index numbers, which define, respectively, the tight upper and lower bounds of the “true” indexes (these bounds can be considered themselves as “true” index numbers).


\[
L_{ij} = W_{ij}^L \quad \text{or} \quad L_{ij} = X_{ij}^L
\]

\[
K_{ij} = W_{ij}^P = \frac{1}{W_{ij}^L} \quad \text{or} \quad K_{ij} = X_{ij}^P = \frac{1}{X_{ij}^L}
\]

and, for \( i \neq j \),

\[
(A2) \quad M_{ij} \equiv \min_{k,l,...,m} L_{nk} L_{kl} ... L_{mj}
\]

which can be called derived (or minimum chained) Laspeyres price or quantity index number

---


\(^{16}\) It may be useful to take note, however, that the fact that the data set satisfies HARP does not exclude that the same data could be also consistent with or generated by some non-conical (non-homothetic) utility (or production) functions if a rational behaviour exists at all.

\(^{17}\) The concepts of aggregate level indexes have been theoretical defined by Keynes (1909), Shephard (1953), Afriat (1953-1956,1972), and Samuelson and Swamy (1974, p. 570).
The computation of the $M$ and $H$ matrices can be based on the application of Edmunds’ (1973) “minimum path” and Bainbridge’s (1978) “power algorithm” to the Laspeyres matrix $L$ for all compared years as adapted by Afriat (1979, 1980b, 1981, 1982). It consists in raising the Laspeyres matrix $L$ to powers $m$ times, with $m$ being the number of the compared observation points, in a modified arithmetic where $+$ means min. In this special arithmetic, in the case of all $M_{ii} \geq 1$, the resulting matrix $M$ remains unchanged if multiplied further by $L$, that is $M \equiv L^m = L^{m+1} = M \cdot L$.

**Proposition 1: Derived LP-inequality condition for aggregation.** The derived Laspeyres ($M_{ij}$) and Paasche ($H_{ij}$) index numbers are the tight bounds of the interval of possible values of the “true” index number and the inequality $M_{ij} \geq H_{ij}$ on the purchaser’s side ($M_{ij} = H_{ij} = K_{ij} = 1$ if HARP conditions are satisfied).

**Proposition 2: Determination of dual price and quantity levels.** If and only if HARP is satisfied, then, in the purchaser’s case, the input price levels $W_i = \theta(w_i)$ associated with the dual input quantity levels $X_i = \phi(x_i)$, such that $W_i \cdot X_i = w_i \cdot x_i$ can be determined, respectively, as solution of the dual inequalities

\[
A4: \quad M_i^W \geq W_i / W_i \\
A5: \quad M_i^X \geq X_i / X_i
\]

whereas, in the supplier’s case, the signs of these inequalities are inverted.

**Proposition 3: The solutions for the dual price and quantity levels are optimal.** HARP conditions also imply,

\[
A6: \quad W_s \cdot X_s \leq w_s \cdot x_s \quad \text{for } s, t = 1, 2, \ldots m
\]

with equality for $s = t$ and reversed inequality sign in the supplier’s case.

**Proposition 8: The “true” price and quantity indexes.** The “true” index numbers can be obtained as ratios between price and quantity level indexes

\[
A7: \quad W_{ij} = W_i / W_j \\
A8: \quad X_{ij} = X_i / X_j
\]
In the discussion of the different approaches to index number theory, van Veelen and van der Weide (2008, p. 1729) conclude: “a challenge for the future is to allow for heterogeneity in index number theory”. Indeed, Afriat’s approach allows the construction of consistent index numbers in the heterogeneity (non-homotheticity) case with data violating HARP. In this case, a suitable correction of the resulting index numbers can be introduced. Allocative inefficiency in both deterministic and stochastic versions has always been viewed to write off the economic meaning of index numbers (see, for example, Mundlak and Razin 1969, p. 169 and Philips, 1983, pp.145-148).

By contrast, Afriat (1967, 1972) introduced an inefficiency index, later called “Afriat’s inefficiency index” by Varian (1990, 1993), as a way to generalize the accounting context thus widening considerably the set of the data that can be rationalized by utility (or production) functions up to a certain degree of efficiency. The starting point for this more general treatment is the recognition that allocative inefficiency is signalled by the violation of the Laspeyres-Paasche inequality condition.

Remark 2: The violation of the LP-inequality condition viewed as lower bound inefficiency in input allocation. In the non-homothetic case characterized by some inequalities $H_{rs}^W > M_{rs}^W$ and $H_{rs}^X > M_{rs}^X$ the violating the derived LP-inequality can be interpreted as lower bound allocative inefficiency in input allocation given by with a utility or technology function rationalizing the data.

The inefficiency effect of misallocation of inputs reducing TFP can be graphically illustrated as in Figure 1.

Figure 1
At equilibrium point $A$ with output $y^0 = f(X^A)$, total cost is given by $W^aX^A$ and total factor productivity level is $\text{TFP}^A = y^0/\lambda^A$. A change in equilibrium due to changes in relative prices, with this level of output, an efficient change would lead the producer at point $C$. At this new equilibrium point, total cost is given by $W^aX^C$ (as $X^A = X^C$).

If the new input allocation is sub-optimal, as for example in $B$, the (in)efficiency cost index is $e = OF/OE$ in the range $0 \leq e \leq 1$. Total cost is now $W^aX^A \frac{1}{e}$. The inefficient use of inputs increases the cost of production by going from $C$ to $B$ violating the LP-inequality condition. With the cost level equal to $OE$, we could produce efficiently $y^1 > y^0$ at point $D$ using the aggregate input quantity $X^D$ as $X^D = X^A \frac{1}{e}$. Therefore, the ratio $y^1/\lambda^0 = f(X^D)/f(X^A)$ is the relative quantity of output loss due to misallocation of inputs occurring by going from $D$ to $B$ along the iso-cost line $DE$. Since, at point $B$, $\text{TFP}^B = y^0/(X^A \frac{1}{e}) = y^0/X^D$, then $\text{TFP}^D = y^1/X^D > \text{TFP}^B$. This is a rationale similar to that motivating the calibrated models recently used to measure the hypothetical loss in the TFP level that would derive from between-firm misallocation of resources. By contrast, the empirical evidence cited above suggests that such a loss is not warranted by the firms’ vertical specialization in China, which would make the possibility of the firms’ output substitution very low. The relevant TFP losses that would arise from misallocation of resources in the Chinese industries are those that occur within, rather than between, the integrated production units. Input misallocation affects the cost efficiency as defined in the following proposition:

**Proposition 4. Cost (in)efficiency** (Afriat 1967, 1972; Varian 1990, 1993). Cost efficiency up to a level $e$, where $0 \leq e \leq 1$ would require

$$W_i \cdot X_i \geq e_i \cdot w_i \cdot x_i = \theta(w_i) \cdot \phi(x_i) \quad (t = 1,2,...,m)$$

which is valid also with any value of cost efficiency lower than $e_0$, whereas the highest possible level of $e_i$ equal to 1 imposes the equality.

Some empirical studies have computed Afriat’s inefficiency index while testing the data for HARP conditions. For example, Feenstra, Ma, Neary, and Rao (2013, p. 1109, fn.15) find that, in the World Bank consumers’ sample of 124 countries in 2005, the value of this index is 0.6620 implying a 33.4% of the budget. The usual procedure confines this non-parametric analysis to the preliminary stage of testing the data for consistency with rationality and homogeneity. In this context, Varian (1990, p. 134) has shown that even a small misallocation waste might be associated with “a bad fit in terms of the usual standard error” of input demands.

In this article, we go further by fully exploiting Afriat’s approach in order to identify the tight first step forward is the correction of the starting Laspeyres matrix for allocative inefficiency. Components (previous practical illustrations are provided in Afriat and Milana, 2009, pp. ch.3). The first step forward is the correction of the starting Laspeyres matrix for allocative inefficiency.

**Proposition 5. Correction for inefficiency in input allocation** (Afriat 1967, 1972; Varian 1990, 1993). If the HARP conditions are not satisfied, then the data can be corrected for inefficiency. A diagonal element $M_{ii} < 1$ tells the inconsistency of the system as it violates the inequality (2a).

**Definition 4:** A critical efficiency parameter $e^*$ can be found for correction of the $L$ matrix. For any element $M_{ii} < 1$, let $d_i$ represent the number of nodes in the cyclical path $i...i$, then

$$e_i = (M_{ii})^{\frac{1}{d_i}}$$

If $M_{ii} \geq 1$, let $e_i$ take the value of 1 while the critical efficiency parameter is determined as

$$e^* = \min \{e_i\}$$
The adjusted Laspeyres matrix is obtained as

\[(A12)\]

\[L'_{ij} = L_{ij} / e^i \quad \text{for} \quad i \neq j\]

and the procedure goes on as before with \(L'\) in place of the original \(L\) in order to compute a new matrix \(M'\) following Definition 1.

The cost-efficiency parameters \(e_i\) are used to compute the efficiency components of the “true” input aggregates, whereas \((1 - e_i)\) represents the cost share of input misallocation. The following propositions define the “true” price and quantity levels of input aggregates.

**Proposition 6: Upper and lower bounds of aggregate input price levels** (Milana 2008). The maximum and minimum aggregate input price levels are the respective elements of the first columns of the matrices \(\mathbf{M}^{w*}\) and \(\mathbf{H}^{w*}\), that is

\[(A13)\]

\[\bar{W}_t^i \equiv \theta(w') = M_{tt}^{w*} \quad \text{for} \quad t = 1, 2, \ldots m\]

\[(A14)\]

\[\underline{W}_t^i \equiv \hat{\theta}(w') = H_{tt}^{w*} = 1/M_{tt}^{w*} \quad \text{for} \quad t = 1, 2, \ldots m\]

with \(\bar{W}_1\) and \(\underline{W}_1\) being set to be equal to 1.

**Proposition 7: Upper and lower bounds of aggregate input quantity levels** (Milana 2008). The upper and lower bounds of aggregate quantity levels are those proportional to the respective elements of the first columns of the matrices \(\mathbf{M}^{x*}\) and \(\mathbf{H}^{x*}\), that is

\[(A15)\]

\[\hat{X}_t = \frac{1}{e_t} \phi(x') = \frac{M_{tt}^{x*}}{w_1(x_1)} \quad \text{for} \quad t = 1, 2, \ldots m\]

\[(A16)\]

\[\tilde{X}_t = \frac{1}{e_t} \hat{\phi}(x') = \frac{H_{tt}^{x*}}{w_1(x_1)} = 1/M_{tt}^{x*} \quad \text{for} \quad t = 1, 2, \ldots m\]

with \(\hat{X}_1\) and \(\tilde{X}_1\) being equal to \(w_1(x_1)\).

In particular cases where it could be necessary to indicate a point estimation rather than the full range of possible values of the “true” index numbers, a geometric average in the spirit of the ideal Fisher number could be provided:

\[(A17)\]

\[\hat{W}_t = \theta(w') = (\bar{W}_t \cdot \underline{W}_t)^{1/2}\]

\[(A18)\]

\[\tilde{W}_t = \frac{1}{e_t} \frac{1}{e_t} \left[\phi(x_1) \cdot \phi(x_1)\right]^{1/2} (w_1(x_1)) = (\hat{X}_t \cdot \tilde{X}_t)^{1/2} (w_1(x_1))\]

for \(t = 1, 2, \ldots m\)

Alternative price level computations have been proposed by Afriat (2008) and adopted in Afriat and Milana (2009), who used the geometric averages of the elements of all the column vectors of the matrices \(\mathbf{M}\) and \(\mathbf{H}\), thus obtaining:

\[(A19)\]

\[\bar{W}_t^A \equiv \theta^A (w') = \left(\prod_{j=1}^m M_{jj}^{w*} \cdot H_{jj}^{w*}\right)^{1/2m} \quad \text{for} \quad t = 1, 2, \ldots m\]

where \(\bar{W}_t^A\), after normalization, may turn out to differ only slightly from \(\bar{W}_t\).
The “true” price and quantity index numbers are obtained as the ratios of the respective “true” price and quantity aggregate levels.

**Remark 5:** Limits of the “true” aggregating input price and quantity index numbers in the general non-homothetic case. The binary “true” index numbers of input prices and quantities obtained as ratios of the respective bounds of aggregate price and quantity levels have the following bounds

\[(A20) \quad \frac{W_{ij}^p}{P_i} \cdot e_{ij} \leq \tilde{W}_{ij} \leq \tilde{W}_{ij} \leq W_{ij}^L / e_{ij} \quad \text{for every } i \geq j\]

\[(A21) \quad L^X_{ij} / e_{ij} \geq \tilde{X}_{ij} \geq \tilde{X}_{ij} \geq K^X_{ij} \cdot e_{ij} \quad \text{for every } i \geq j\]

with \(e_{rs}\) representing the index of cost efficiency index while the following equality always holds:

\[(A22) \quad \frac{W_{ij}^L}{W_{ij}^P} = X_{ij}^L = \tilde{W}_{ij} \cdot \tilde{X}_{ij} = \tilde{W}_{ij} \cdot \tilde{X}_{ij} = W_{ij}^P \cdot X_{ij}^L \quad \text{for every } i, j = 1, 2, \ldots m\]

**Remark 6:** All Fisher’s tests are satisfied. As shown by Afriat (1972) and Samuelson and Swamy (1974), all the “true” index numbers that are constructed as ratios between aggregate level indexes, satisfy all Fisher’s requirements including transitivity.
References


Bank of Italy. 2008. Financial accounts: History, Methods, the Case of Italy and International Comparisons. Rome.


