# Evaluating and improving international assistance programmes: Examples from Mongolia's transition experience

Schouwstra, M.C.

**Publication date**
2013

# 4.
# Total Factor Productivity and the Mongolian Transition

*Marije C. Schouwstra and Antonio G. Chessa*
*University of Amsterdam, The Netherlands*

## *4.1 Introduction*

The objective of growth accounting is to break down the growth rate of aggregate output into contributions from the growth of inputs, usually capital and labour, and the growth of technology (Barro and Sala-i-Martin, 1995). The analysis starts from the Cobb-Douglas function. Total Factor Productivity is the residual of the Cobb-Douglas function or the unexplained portion of GDP growth (Fajnzylber and Lederman, 1999; Hulten, 2000). It is often used on the macroeconomic level as an indicator of efficiency of a country[81]. IMF, World Bank and many other economists use the concept to compare various countries with each other or to compare the performance of a country over different years. In this chapter, several mathematical properties of the concept of TFP are examined, which will be used for explaining the development of the yearly changes in TFP for Mongolia in particular.

In most transition economies, Total Factor Productivity was negative in the last decade before the transition and started to increase after (quite a few) years of transition (Kaser, 2004). All transition economies (except China and Vietnam) went through a phase of severe recession just after the transition –the length of which varied considerably amongst the various countries. In Mongolia enterprises tried to survive this period by cutting investment, by temporarily interrupting operations (ADB, 1996), shortening working hours or by not paying their employees. Thus, in the first instance they retained their labour. Despite these measures, many enterprises did have to lay-off workers and had to close down partially or completely after a few years after all. At the same time the informal economy had been growing considerably and absorbed many of those who lost their jobs in the formal sector. The international organisations had advised the Mongolian government on a rigorous privatization and liberalization of their economy. When enterprises did close down and TFP did indeed start to increase after the first few years of

[81] In the 1970s this factor was called the contribution of technical progress to the growth of a country.

the transition, they – and many other authors – attributed these gains in TFP to gains in efficiency due to the structural reform measures carried out in order to establish a market economy (see, amongst others, Cheng, 2003; IMF, 2003). The remarkable thing was however that the enterprises that closed down in Mongolia were often not the least promising. Many enterprises that could possibly be viable and had potential closed down whereas some of the (often large) non-viable enterprises remained open, often with the help of the government. In 2002 the World Bank was considering whether to support the reopening of production facilities in Mongolia that had closed in the wake of the transition, in order to combat unemployment in various regions, which they would certainly not have considered if those facilities had not been potentially viable. This raised the question of the meaning of the development of TFP and whether the changes in TFP were really attributable to increases in efficiency or whether other factors – not only temporary phenomena such as the retention of labour during a recession period but phenomena such as the closure of a large number of enterprises in an economy – were significant in explaining, at least part of, the development of TFP since the transition.

In order to investigate this proposition, this chapter looks into the mathematical properties of the Cobb-Douglas function to see whether these have any explanatory value with regard to the development of Total Factor Productivity for Mongolia in the first decade of its transition. In Section 4.2 the definitions and properties of TFP will be outlined and in Section 4.3 Total Factor Productivity will be analysed for Mongolia. In Section 4.3.1 the Cobb-Douglas function is described for Mongolia. In the next section the development of TFP for Mongolia is looked into and in Section 4.3.3 the question of data quality is addressed. In Section 4.3.4 a sensitivity analysis is carried out, both for the parameters of the Cobb-Douglas function and for possible errors in the data. To test the hypothesis that randomly closing enterprises – instead of closing only the least efficient enterprises – has an influence on the development of TFP comparable to the development of TFP such as is witnessed in Mongolia, we have simulated a closed economy with 1000 enterprises in Section 4.4. In the simulations a certain number of enterprises is closed under such restrictions as are observed in Mongolia. The advantage of simulations is that the usual assumptions with regard to TFP, such as marginal pricing or the closure of only the least efficient enterprises (phenomena that are not always observed in the real world) need not be imposed. Finally, in Section 5 the conclusions are presented.

## *4.2 Definitions and properties of TFP*

In this section, we will define TFP and derive some of its properties that may be helpful in gaining more insight into the possible reasons behind the increase of TFP in Mongolia after several years of transition. We will apply the formal results of our derivations in Section 4.3 for the specific case of Mongolia's economy.

The starting point of our analysis is the representation of a closed macroeconomic system as a triplet (*Y*, *K*, *L*), the components of which respectively denote Gross Domestic Product (GDP), capital stock and labour. As in other, previous studies we also assume that GDP satisfies the Cobb-Douglas production function

$$Y = AK^{\alpha}L^{1-\alpha}, \tag{1}$$

where $0 \leq \alpha \leq 1$ and $A$ is a nonnegative constant, so that gross domestic product $Y$ is increasing in both capital $K$ and labour $L$. The parameter $\alpha$ is referred to as 'the income share of capital'.

The choice of the Cobb-Douglas relation between GDP, capital and labour can be justified from a measurement theoretic point of view (Roberts, 1979). This function satisfies the property that comparisons of GDP among different companies, or over different years, are invariant under scale transformations of the inputs capital and labour. Capital stock is expressed in monetary units and labour in terms of the number of persons employed[82]. The two inputs in (1) are valued on ratio scales $i = 1, 2$, with values on scale $i$ denoted by $x_i$. These scales are completely described by the class of similarity transformations $\phi_i : \mathrm{R} \to \mathrm{R}$, which are given by $\phi_i(x_i) = a_i x_i$, $a_i > 0$, $i = 1, 2$, where the scale constants $a_1$ and $a_2$ may differ. Similarity transformations define the class of admissible transformations for ratio scales.

The rationale behind Theorem 1 below is that statements about GDP and TFP must be invariant under all ratio scale transformations of the input variables, for instance, when converting capital from US Dollars into Euros. To illustrate matters we consider the following situation, which may serve to set ideas for the proof of the theorem. Suppose that GDP has a maximum on some set $G$ of capital-labour pairs $(x_1, x_2)$, say $G = \{(x_1, x_2) \in \mathrm{R}_+^2 : c_1 x_1 + c_2 x_2 \leq c\}$, where $c_1$, $c_2$ and $c$ are nonnegative constants. Since GDP is increasing in both variables, it has a maximum, $(m_1, m_2)$ say, on the set $G' = \{(x_1, x_2) \in \mathrm{R}_+^2 : c_1 x_1 + c_2 x_2 = c\}$. This optimality must be preserved under all ratio scale transformations of the inputs (see Figure 4-1).

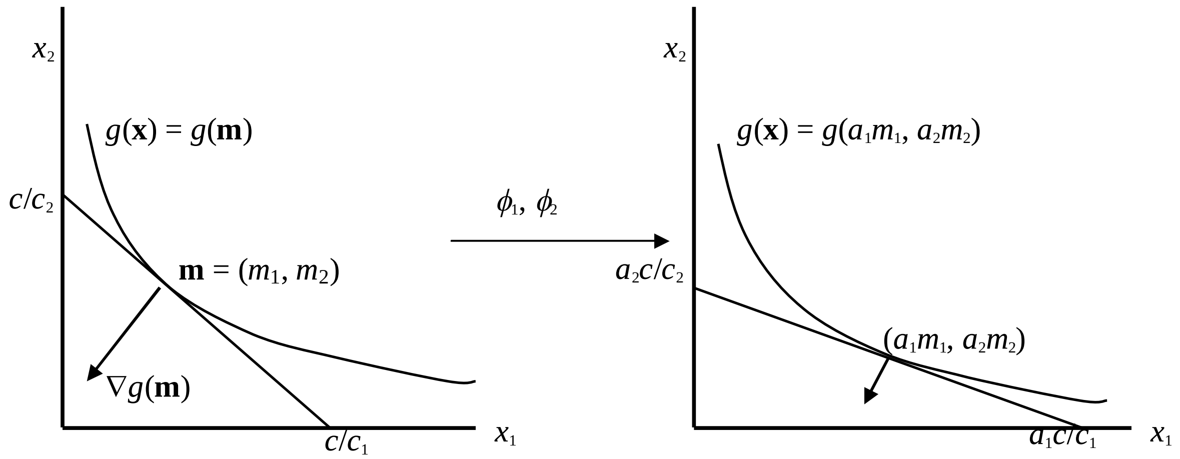


**Figure 4-1. Relation between the gradients of g in optima m = (m1, m2) and (a1m1, a2m2) and their invariance under similarity transformations of two ratio scales.**

[82] Often some arithmetic calculation is done with the number of persons employed to obtain a measure of "human capital", e.g. in Cheng (2003) this number is multiplied by the number of years of schooling.

Here we only state Theorem 1; the proof is given in Appendix A4.

THEOREM 1. *A function g on* $\mathbb{R}_+^2$, *with partial derivatives existing in all* $x_1, x_2 > 0$, *where, in all applications,* $g(x_1, x_2)$ *denotes GDP, and* $x_1$ *and* $x_2$ *denote capital and labour inputs measured on ratio scales, is uniquely described by*

$$g(x_1, x_2) = \Phi\left(\ln\left(x_1^{\,a} x_2^{\,b}\right)\right), \tag{2}$$

*where a, b are real-valued constants and* $\Phi$ *is an exponential function.*

Function (2) is a generalised version of the Cobb-Douglas production function

$$g(x_1, x_2) = A x_1^{\,\alpha} x_2^{\,1-\alpha}, \tag{3}$$

where $0 \le \alpha \le 1$ and $A$ is a nonnegative constant, which is the function used by Cheng (2003). One of the assumptions in (3) is that of constant returns to scale, which holds when $a + b = 1$ in (2) (Chiang, 1984). In the rest of this chapter, we will work with (3). This function has a number of implications, which are not restricted to the applications described in this chapter, but are interesting in their own right as well:

*Decision analysis.* The result of Theorem 1 is important for decision-making problems where alternatives are evaluated and ranked according to two criteria. If value judgments of the alternatives are expressed on ratio scales for both criteria, then Theorem 1 says that the ranking of alternatives is determined by combining the values judgments for the two criteria by a function $g$, for every alternative. The parameter $\alpha$ is in fact the relative weight of criterion 1.

*Production functions*. The Cobb-Douglas function is well known in growth accounting, and other meanings of $g$ and of the parameter $\alpha$ derive from this field. Function (3) may represent a production function with two inputs, in which $\alpha$ is equal to the output elasticity with respect to input $x_1$, that is:

$$\frac{\partial g / \partial x_1}{g / x_1} = \alpha, \tag{4}$$

which represents a relative marginal change in the output $g$ due to a relative marginal change in the input $x_1$. Within the context of this chapter, the parameter $\alpha$ can thus be interpreted as the elasticity of GDP with respect to capital. A small value of $\alpha$ means that GDP is mainly determined by labour. Another appealing notion that involves $\alpha$ is the technical elasticity of substitution, which can be written as $(1 - \alpha)/\alpha$ (or by its inverse). This ratio denotes the relative amount of labour that can be substituted by capital in order to yield the same GDP.

*Total Factor Productivity*. The factor $A$ is a scaling term that relates GDP to capital and labour. This is the term that is known as *Total Factor Productivity* or the unexplained portion of GDP (Fajnzylber and Lederman, 1999). It is being interpreted as an indicator of the efficiency of a country's economy: By rewriting GDP as $Y^{\alpha}Y^{1-\alpha}$ and by dividing it by the right-hand side of (1), without involving $A$, we obtain the expression

$$A = \left(\frac{Y}{K}\right)^{\alpha} \left(\frac{Y}{L}\right)^{1-\alpha}, \tag{5}$$

where $Y/K$ and $Y/L$ denote GDP per unit of capital and labour (*capital-productivity* and *labour-productivity* respectively). By the uniqueness of the result of Theorem 1, TFP receives a strict meaning according to (5).

We will now derive some properties of TFP that may give some insight into the possible causes of the TFP decrease before and directly after the transition to a market economy of Mongolia and its increase after the first few years of transition. For this purpose, we take a closed economy with two enterprises as our starting point. We subdivide this economy, which we represent by the triplet $(Y, K, L)$ (production, capital and labour respectively), into two subsystems $(Y_1, K_1, L_1)$ and $(Y_2, K_2, L_2)$, such that $Y = Y_1 + Y_2$, $K = K_1 + K_2$ and $L = L_1 + L_2$. Furthermore, the two subsystems also satisfy the Cobb-Douglas function in the sense that

$$Y_i = A_i K_i^{\alpha} L_i^{1-\alpha}, \tag{6}$$

where $A_i$ denotes the TFP of enterprise $i = 1, 2$. With regard to (6), we make the following two assumptions: (1) $0 \leq \alpha \leq 1$, and (2) $\alpha$ is the same for all subsystems $I$, and for all sets of subsystems. The following theorem lists a number of properties of TFP, some of which are generalised to economies with $n$ subsystems (Theorem 3).

THEOREM 2. *Let* $\kappa = K_1/(K_1 + K_2)$, $\lambda = L_1/(L_1 + L_2)$ *and* $y = Y_1/(Y_1 + Y_2)$. *The above macroeconomic model, which is described by functions (1) and (6), has the following properties*:

(i) *If* $Y_1/L_1 > Y_2/L_2$ *and* $Y_1/K_1 > Y_2/K_2$, *then* $A_1 > A > A_2$ *always holds, for every* $\alpha \in [0, 1]$;

(ii) *If* $Y_1/L_1 > Y_2/L_2$ *or* $Y_1/K_1 > Y_2/K_2$, *but not both, then an economy exists, such that* $A_1 < A < A_2$;

(iii) *Both* $A_1 > A$ *and* $A_2 > A$ *hold if, and only if*

$$\kappa^{\alpha}\lambda^{1-\alpha} < y < 1-(1-\kappa)^{\alpha}(1-\lambda)^{1-\alpha}, \quad \lambda \neq \kappa,\ 0 < \alpha < 1. \tag{7}$$

(iv) *There is no economy such that both* $A_1 < A$ *and* $A_2 < A$.

PROOF.

*Proof of property (i).* The inequality $Y_1/L_1 > Y_2/L_2$, which we rewrite as $Y_1/Y_2 > L_1/L_2$, for nonzero $Y_2$, implies that $Y_1/(Y_1 + Y_2) > L_1/(L_1 + L_2)$, and therefore also $Y_1/L_1 > Y/L$, which also holds for zero $Y_2$. Along similar lines it can be shown that $Y/L > Y_2/L_2$, so that $Y_1/L_1 > Y/L > Y_2/L_2$. The same inequality can be proved for capital, that is, $Y_1/K_1 > Y/K > Y_2/K_2$. The inequality $A_1 > A > A_2$ now directly follows from expression (5) for TFP, which holds for all $\alpha \in [0, 1]$.

*Proof of property (ii).* Consider the inequality $Y_1/L_1 > Y_2/L_2$. On the basis of the stated condition on capital we can choose $K_1$ and $K_2$ such that $Y/K > Y_1/K_1$. By choosing $\alpha$ 'large enough' (i.e., close to, or equal to 1), we obtain $A > A_1$ by using expression (5). The existence of situations where $A < A_2$ can be proved in the same way. The proof follows a similar reasoning for situations where $Y_1/K_1 > Y_2/K_2$ and will therefore be omitted.

*Proof of property (iii).* By definition of TFP, $A_1 > A$ and $A_2 > A$ both hold if, and only if (iff)

$$\frac{Y_i}{K_i^{\alpha} L_i^{1-\alpha}} > \frac{Y_1 + Y_2}{(K_1 + K_2)^{\alpha}(L_1 + L_2)^{1-\alpha}},$$

for $i = 1, 2$, which we rewrite as

$$\frac{Y_i}{Y_1 + Y_2} > \left(\frac{K_i}{K_1 + K_2}\right)^{\alpha} \left(\frac{L_i}{L_1 + L_2}\right)^{1-\alpha}, \tag{8}$$

$i = 1, 2$. For $i = 2$, it follows from (8) that

$$\frac{Y_2}{Y_1 + Y_2} = 1 - \frac{Y_1}{Y_1 + Y_2} > \left(\frac{K_2}{K_1 + K_2}\right)^{\alpha} \left(\frac{L_2}{L_1 + L_2}\right)^{1-\alpha}, \tag{9}$$

so that

$$\frac{Y_1}{Y_1 + Y_2} < 1 - \left(\frac{K_2}{K_1 + K_2}\right)^{\alpha} \left(\frac{L_2}{L_1 + L_2}\right)^{1-\alpha}, \tag{10}$$

or, in the shorthand notation introduced in the theorem,

$$y < 1 - (1-\kappa)^{\alpha}(1-\lambda)^{1-\alpha}. \tag{11}$$

From (8) we have $y > \kappa^{\alpha}\lambda^{1-\alpha}$ for $i = 1$, which gives, together with (11):

$$\kappa^{\alpha}\lambda^{1-\alpha} < y < 1 - (1-\kappa)^{\alpha}(1-\lambda)^{1-\alpha}. \tag{12}$$

Both $A_1 > A$ and $A_2 > A$ hold iff values of $\kappa$, $\lambda$ and $\alpha$ exist, such that the difference between the right-hand and left-hand side of (12) is greater than zero, that is, iff

$$1-(1-\kappa)^{\alpha}(1-\lambda)^{1-\alpha}-\kappa^{\alpha}\lambda^{1-\alpha}>0. \tag{13}$$

We denote the left-hand side of (13) as the function $f(\lambda, \kappa)$. By setting the partial derivatives of $f$ with respect to $\lambda$ and $\kappa$ equal to zero, we obtain the relation $\lambda = \kappa$, which is a minimum of $f$. Since $f(\lambda, \lambda) = 0$ for every $\lambda$, $\lambda$ cannot be equal to $\kappa$ in order for (13) to hold. The same holds for $\alpha = 0$ and $\alpha = 1$. Inequality (13) holds for all other values of $\kappa$, $\lambda$ and $\alpha$.

*Proof of property (iv).* By following the same lines of the proof of property (iii), it follows that $1 - (1 - \kappa)^{\alpha}(1 - \lambda)^{1-\alpha} < y < \kappa^{\alpha}\lambda^{1-\alpha}$ when both $A_1 < A$ and $A_2 < A$. From (13) and the analysis immediately following that expression, it follows that the aforementioned inequality never holds, which proves (iv).

Notice that properties (i) and (iv) do not depend on the value of $\alpha$. This is a very important result for property (i) in particular, because it enables us to draw conclusions about TFP-growth without knowing the value of $\alpha$, by simply checking whether the conditions stated in property (i) hold for time series of GDP, capital and labour. We will illustrate the use of these conditions in Section 4.3 where we analyse time series for Mongolia.

The above properties can be translated into implications for an economy and its TFP when enterprises are closed, without reallocating its capital and labour elsewhere, or when new capital and labour are invested. From property (i) it follows that TFP always increases from $A$ to $A_1$ when an enterprise with the lowest capital and labour productivities is closed (i.e., subsystem 2), if such an enterprise exists. Property (i) also tells us that TFP always increases from $A_2$ to $A$ when new amounts of capital $K_1$ and labour $L_1$ are invested that give rise to a contribution $Y_1$ to the existing production $Y_2$ that is more capital- and labour-productive (i.e., $Y_1/L_1 > Y_2/L_2$ and $Y_1/K_1 > Y_2/K_2$). It is important to note that these behaviours of TFP hold irrespective of the value of $\alpha$.

Property (ii) says that these two situations do not necessarily arise when there are no enterprises, of which the productivities satisfy the conditions stated in property (i). Two situations occur under property (ii): (1) closing an enterprise (i.e., subsystem 2) with the smallest labour productivity (capital productivity), but which has not the smallest capital productivity (labour productivity), (2) new capital and labour resources are invested with one of the two resources having a larger productivity and the other resource a lower productivity than that of the existing enterprises. In case (1), TFP may decrease from $A$ to $A_1$, while TFP may decrease from $A_2$ to $A$ in case (2).

Property (iii) shows that the TFP of single enterprises may be greater than the TFP of the aggregate economy. This arises when production, capital and labour in the two subsystems satisfy condition (7). If this condition is satisfied, then the closing of an arbitrary enterprise will lead to an increase of TFP. Similarly, if new capital and labour are invested such that condition (7) holds, then TFP will decrease.

A more detailed investigation of property (iii) also shows that the distribution of capital and labour over enterprises affects TFP under reforms. Let capital and labour share be related according to a ratio $c$ as $\kappa = c\lambda$. Then inequality (7) becomes

$$c^{\alpha}\lambda < y < 1-(1-c\lambda)^{\alpha}(1-\lambda)^{1-\alpha}. \tag{14}$$

An interesting aspect of this inequality is its behaviour when varying $c$. For example, when $c$ goes to zero we obtain $0 < y < 1 - (1 - \lambda)^{1-\alpha}$. Letting $c$ go to zero implies that capital in subsystem 2 is much greater than in subsystem 1. If we have the reverse situation for labour share, so that $\lambda$ goes to 1, then we obtain, in the limit, $0 < y < 1$. This illustrates that conclusions about TFP will depend less on $Y_i$ and $\alpha$ when capital and labour are distributed in a less balanced and less uniform way over enterprises. The result stated in property (iii) implies for instance that TFP is likely to increase when closing an arbitrary enterprise, or ensemble of enterprises, which have a very large capital share and a very small labour share, or vice versa. A similar reasoning applies when making new investments of this kind, in which case TFP is likely to decrease.

Property (iii) is interesting from a 'managerial' point of view, in the sense that the three terms of inequality (7) either depend only on the input variables or only on $Y$. By making decisions regarding the inputs, the implications for TFP can be controlled to a certain extent, even without knowledge of $Y$ and $\alpha$.

Property (iv) says that it is impossible to have an economy with a TFP that is greater than the TFP of two single enterprises or subsystems. In other words, an economy as a whole cannot be more efficient than all its subsystems. Property (iv) also implies that investments of arbitrary amounts of capital and labour will not necessarily lead to a higher TFP.

Theorem 2 can be generalised to situations with $n$ subsystems or enterprises. Below we will give two generalisations of property (i). We also formulate an additional property of TFP for $n$ enterprises, which is a fundamental result in itself, and which will be useful in the interpretation of certain results in Section 4.

THEOREM 3.

(i) *Let I and J denote disjoint nonempty sets of enterprises and let $A_I$, $A_J$ and $A_{I\cup J}$ denote the TFP of the ensembles of enterprises in I, in J and in the union of I and J respectively. The inequality $A_I > A_{I\cup J} > A_J$ holds, for every $\alpha \in [0, 1]$, in the following situations:*

(a) *if* $\dfrac{\sum_{i\in I} Y_i}{\sum_{i\in I} K_i} > \dfrac{\sum_{j\in J} Y_j}{\sum_{j\in J} K_j}$ *and* $\dfrac{\sum_{i\in I} Y_i}{\sum_{i\in I} L_i} > \dfrac{\sum_{j\in J} Y_j}{\sum_{j\in J} L_j}$,

(b) *if* $Y_i/K_i > Y_j/K_j$ *and* $Y_i/L_i > Y_j/L_j$*, for all* $i \in I$ *and all* $j \in J$*.*

(ii) *Let Y, K and L denote the sums of production* $Y_i$*, of capital stock* $K_i$ *and of labour* $L_i$ *over enterprises* $i = 1, 2, \ldots, n$*, and let* $\kappa_i = K_i/K$ *and* $\lambda_i = L_i/L$ *for all i. The TFP* $A_1, \ldots, A_n$ *of the individual enterprises and the TFP A of the aggregate economy satisfy the following relation:*

$$A = \sum_{i=1}^{n} \kappa_i^{\alpha} \lambda_i^{1-\alpha} A_i. \tag{15}$$

PROOF.

*Proof of property (i)*. The first result of (i) follows directly from Theorem 2.(i). In order to prove the second result, let $n \in J$ denote an enterprise with the largest capital-productivity among the enterprises in $J$. Then it follows that

$$\frac{\sum_{j\in J} Y_j}{\sum_{j\in J} K_j} \le \frac{\sum_{j\in J} K_j Y_n / K_n}{\sum_{j\in J} K_j} = \frac{Y_n}{K_n}, \tag{16}$$

so that the capital-productivity of enterprise $n$ is also at least as great as the capital-productivity of ensemble $J$. Since $Y_i/K_i > Y_n/K_n$, for all $i \in I$, it follows that

$$\frac{\sum_{i\in I} Y_i}{\sum_{i\in I} K_i} > \frac{Y_n}{K_n} \ge \frac{\sum_{j\in J} Y_j}{\sum_{j\in J} K_j}, \tag{17}$$

where the second inequality arises from (16). Inequality (17) shows that the capital-productivity of ensemble $I$ is greater than that of ensemble $J$. The same can be proved for the labour-productivities of $I$ and $J$ by following the same reasoning. The first result of (i) thus implies that $A_I > A_{I\cup J} > A_J$, irrespective of the value of $\alpha$.

*Proof of property (ii)*. This result follows by rewriting expression (5) for TFP, which we first write as

$$A = \sum_{i=1}^{n} \frac{Y_i}{K^{\alpha} L^{1-\alpha}}. \tag{18}$$

We finalise the proof by multiplying the ratio under the summation by ratios equal to 1 in the following way:

$$A = \sum_{i=1}^{n} \frac{K_i^{\alpha} L_i^{1-\alpha}}{K_i^{\alpha} L_i^{1-\alpha}} \frac{Y_i}{K^{\alpha} L^{1-\alpha}} = \sum_{i=1}^{n} \kappa_i^{\alpha} \lambda_i^{1-\alpha} \frac{Y_i}{K_i^{\alpha} L_i^{1-\alpha}} = \sum_{i=1}^{n} \kappa_i^{\alpha} \lambda_i^{1-\alpha} A_i. \tag{19}$$

This holds when $K_i$ and $L_i$ are positive for all $i$. If $K_i = L_i = 0$ for some $i$, then TFP $A_i$ is not defined. If $K_i = 0$ and $L_i > 0$ for some $i$, then $A_i$ is only defined when $\alpha = 0$, and expression (19) can then be proved in the same way as above by taking $\alpha = 0$. The same holds when $L_i = 0$ and $K_i > 0$ for some $i$, in which case $\alpha = 1$ in order for $A_i$ to be defined. This completes the proof.

Property (i) of this theorem says that TFP will increase when additional capital and labour are invested in new or existing enterprises, which, in aggregate terms, have overall capital- and labour-productivity that exceed the respective productivities of the existing enterprises taken together. A similar implication follows when enterprises with the smallest overall capital- and labour-productivity are closed. Property (i)-(b) is a special case, as it gives information about the capital- and labour productivities of every single investment in capital and labour.

Property (ii) of Theorem 3 gives the exact aggregation rule for combining the TFP of the individual enterprises in order to derive the TFP for the overall economy. It shows that overall TFP is an arithmetic average, which is weighted however by a product involving the capital and labour shares of each enterprise. The shares themselves are weighted by $\alpha$ and $1 - \alpha$ respectively. So, overall TFP is not simply the average of the individual TFP-values. Enterprises with a small capital or labour share are likely to give a small contribution to overall TFP (i.e., if their production is not exceedingly high).

## *4.3 Analysis of TFP for Mongolia*

### 4.3.1 The Cobb-Douglas function

We rewrite the Cobb-Douglas function (1) for Gross Domestic Product in year $t$ ($Y_t$) as

$$Y_t = A_t K_t^{\alpha} (q_t L_t)^{1-\alpha}, \tag{20}$$

where $K_t$ is the capital stock, $q_t$ is a human capital index, $L_t$ is labour input and $A_t$ is TFP in year $t$. Capital stock is calculated with the conventional perpetual inventory method:

$$K_{t+1} = I_{t+1} + (1-\delta)K_t, \tag{21}$$

where $I_t$ is the level of real investment in year $t$ and $\delta$ is the rate of depreciation of the capital stock. Usually, capital stock in each successive year is calculated recursively according to (21) by making use of time series data for investment $\{I_t\}$. As a starting point for calculations, capital stock is assumed to be zero in a certain year (if data are not available such as is the case for Mongolia). The level of the depreciation rate is most often assumed to have a value between the extreme values of 0.04 and 0.10. In his Mongolia-study, Cheng (2003) assumed that $\delta = 0.06$.[83] If we denote initial capital stock by $I_0$ for notational convenience, then equality (21) leads to the following expression:

$$K_t = \sum_{i=0}^{t} (1-\delta)^{t-i} I_i. \tag{22}$$

[83] Cheng assumed $\delta = 0.06$ but did not give any arguments why he chose 6% and not any other percentage as depreciation rate. In his calculations of capital stock he used the time series data for investment $\{I_t\}$ since 1980 and furthermore assumed capital stock to be zero in 1959. He performed a sensitivity analysis to test the robustness of his “1959”-assumption.

Due to a lack of official data, especially the lack of data on earnings by educational status for Mongolia, we cannot use the conventional measure of human capital, which is calculated by summing over the number of workers with different educational levels, weighted by their earnings. Instead we use – following Cheng (2003) – total employment as a proxy for labour input, and express the human capital index in terms of the average number of years of schooling of the Mongolian population aged 15 or above. Data are available for two years only: the average number of years of schooling in Mongolia is equal to 7.5 in 1989 and is equal to 8.5 in 1998. Human capital index values are derived – again following Cheng (2003) – by assuming education to be a linear function of time. If we fix $t$ = 0 at the year 1980, that is, at the first year in the time series, we can express human capital in year $t$ in terms of the two data points as follows:

$$q_t = (q_{18} - q_9)t/9 + 2q_9 - q_{18}, \tag{23}$$

where $q_9$ and $q_{18}$ are the human capital values for 1989 and 1998, respectively.

Together with the expression for TFP at time or year $t$, which is equal to

$$A_t = \left(\frac{Y_t}{K_t}\right)^{\alpha}\left(\frac{Y_t}{q_t L_t}\right)^{1-\alpha}, \tag{24}$$

the model that we will analyze in subsequent sections is given by the expressions (22)-(24). Notice that we have formulated these expressions of the model such that this system is completely described by the parameters $\alpha$ and $\delta$, the time variable $t$ and the data variables: GDP $Y_t$, total employment $L_t$, the average years of schooling in 1989 and 1998 ($q_9$ and $q_{18}$), real investment $I_t$ and the capital stock $I_0$ in 1980.

### 4.3.2 Development of TFP

We will start this section by analyzing the growth accounting data with regard to Mongolia as given by Cheng (2003, Appendix Table 2) according to the models described in Sections 2 and 3.1. To this end, we calculated the values of capital- and labour-productivity for the period 1980-2001, which we also did for the 'incremental values'[84] of GDP, capital and labour (the last of which was multiplied by the years of schooling to create a measure for "human capital") between successive years. These values are given in Table 4-1. For instance, labour productivity for the incremental values (Δ Lab. Prod., see Table 4-1) in 1986 is the ratio of GDP in 1986 minus GDP in 1985 and the difference between the values for labour-times-years of schooling in these two years. The productivities for the incremental values may be negative. This occurs, for instance, when GDP increases while capital stock decreases.

[84] The term 'incremental value' is introduced here to denote the difference between the values of a variable in two successive years. More precisely, the incremental value of e.g. GDP in the year 1985 is equal to GDP in 1985 minus GDP in 1984.

**Table 4-1. GDP, capital- and labour-productivity, and TFP for Mongolia from 1980 until 2001. GDP and capital are expressed in billions of tugriks (constant 1995 prices) and labour in thousands of persons times schooling years.**

| Year | GDP | Cap. Prod. | Lab. Prod. | Δ Cap Prod | Δ Lab Prod | TFP (x 10) $\alpha = 1/3$ | $\alpha = 1/2$ | $\alpha = 0.69$ |
|---|---|---|---|---|---|---|---|---|
| 1980 | 376.1 | 0.318 | 0.109 | | | 1.555 | 1.859 | 2.279 |
| 1981 | 407.4 | 0.302 | 0.117 | 0.191 | 2.336 | 1.609 | 1.883 | 2.254 |
| 1982 | 441.1 | 0.291 | 0.122 | 0.199 | 0.227 | 1.629 | 1.882 | 2.221 |
| 1983 | 466.9 | 0.286 | 0.125 | 0.219 | 0.202 | 1.643 | 1.887 | 2.208 |
| 1984 | 494.9 | 0.279 | 0.128 | 0.202 | 0.266 | 1.664 | 1.893 | 2.194 |
| 1985 | 525.7 | 0.273 | 0.126 | 0.201 | 0.092 | 1.627 | 1.851 | 2.145 |
| 1986 | 575.0 | 0.273 | 0.124 | 0.276 | 0.111 | 1.615 | 1.842 | 2.139 |
| 1987 | 594.8 | 0.263 | 0.122 | 0.130 | 0.087 | 1.581 | 1.796 | 2.077 |
| 1988 | 625.2 | 0.261 | 0.114 | 0.217 | 0.047 | 1.499 | 1.721 | 2.016 |
| 1989 | 651.5 | 0.256 | 0.114 | 0.178 | 0.114 | 1.490 | 1.706 | 1.990 |
| 1990 | 635.1 | 0.244 | 0.107 | -0.311 | -0.073 | 1.406 | 1.614 | 1.890 |
| 1991 | 576.4 | 0.227 | 0.094 | 0.996 | -0.342 | 1.262 | 1.461 | 1.727 |
| 1992 | 521.6 | 0.214 | 0.082 | 0.562 | -0.228 | 1.127 | 1.323 | 1.587 |
| 1993 | 505.9 | 0.209 | 0.083 | 0.648 | 0.064 | 1.126 | 1.315 | 1.569 |
| 1994 | 517.6 | 0.219 | 0.084 | -0.223 | 0.375 | 1.157 | 1.357 | 1.627 |
| 1995 | 550.3 | 0.238 | 0.087 | -0.648 | 0.234 | 1.220 | 1.442 | 1.743 |
| 1996 | 563.2 | 0.248 | 0.088 | -0.271 | 0.138 | 1.245 | 1.480 | 1.802 |
| 1997 | 585.7 | 0.261 | 0.091 | -0.913 | 0.575 | 1.294 | 1.543 | 1.884 |
| 1998 | 606.4 | 0.271 | 0.090 | -4.086 | 0.067 | 1.300 | 1.562 | 1.926 |
| 1999 | 625.9 | 0.278 | 0.088 | 1.580 | 0.057 | 1.296 | 1.568 | 1.950 |
| 2000 | 632.5 | 0.279 | 0.089 | 0.351 | 0.161 | 1.301 | 1.574 | 1.956 |
| 2001 | 639.7 | 0.277 | 0.086 | 0.181 | 0.025 | 1.274 | 1.547 | 1.930 |

Source: GDP and labour are taken from Cheng (2003, Appendix Table 2, p.18); Capital has been calculated by the authors.

Cap. Prod. = GDP/Capital

Lab. Prod. = GDP/(Labour x Years of schooling)

Δ Cap Prod in year $t$ = (GDP in year $t$ – GDP in year $t$ – 1)/(capital in year $t$ – capital in year $t$ – 1)

Δ Lab Prod = (GDP in yr $t$ – GDP in yr $t$ – 1)/(lab. x schooling yrs in yr $t$ – lab. x schooling yrs in yr $t$ – 1)

By associating one of the triplets ($Y_1$, $K_1$, $L_1$) and ($Y_2$, $K_2$, $L_2$) with GDP, capital and labour in a year and the other triplet with the incremental values with respect to the next year, we can derive implications for TFP by applying Theorem 2. We illustrate this by the following two examples.

- *TFP in 2000*. We compare labour-productivity for the incremental values of GDP and labour in 2000 (Δ Lab. Prod. = 0.161, see Table 4-1.) with labour-productivity for the absolute values of GDP and labour in 1999 (Lab. Prod. = 0.088). We do the same for capital (Δ Cap. Prod. in 2000 = 0.351; Cap. Prod. in 1999 = 0.278). Let us denote capital- and labour-productivity in 1999 by $Y_2/K_2$ and $Y_2/L_2$, and the productivities for the above incremental values in 2000 by $Y_1/K_1$ and $Y_1/L_1$, by making use of the notation in Theorem 2. Then we have that $Y_1/L_1 > Y_2/L_2$ and $Y_1/K_1 > Y_2/K_2$, so that property (i) of Theorem 2 applies. This property tells us that TFP increases from $A_2$ to $A$, which corresponds with TFP in 1999 and 2000 respectively, and that the increase, on an ordinal scale, does not depend on the value of $\alpha$. The term 'ordinal' refers to a statement that is merely based on the ranking of TFP-values in different years. The numerical values of TFP themselves do change under different $\alpha$-values, as can be seen in Table 4-1.[85]

- *TFP in 2001*. In this case, the values of capital- and labour-productivity for the incremental values of GDP, capital and labour in 2001 are both smaller than capital- and labour-productivity in 2000. From property (i) of Theorem 2 it follows that TFP in 2001 is smaller than TFP in 2000.

These examples show that Theorem 2 tells us that new investments in capital and labour positively affected TFP in 2000, while the contrary applies to 2001. Theorem 3 can be used to give more specific, possible explanations of these developments by considering an economy with $n$ enterprises. For example, the investments made for capital and labour in new or existing enterprises in 2000 may each have greater capital- and labour-productivity than the capital- and labour-productivities of the existing enterprises (Theorem 3, property (i)-(b)), or the capital- and labour-productivities of the ensemble of new investments exceed the productivities of capital and labour in 1999 (Theorem 3, property (i)-(a)).

When we apply such analyses to the period 1980-2001, we come to the following characterisation of Mongolia's macroeconomic situation:

- 1980 – 1984: Labour-productivity increases, while capital-productivity decreases. This is a situation that falls under property (ii) of Theorem 2, so that TFP may either increase or decrease, depending on the value of $\alpha$. A value of $\alpha$ close to 0 attaches most weight to labour, so that TFP would increase. A value of $\alpha$ close to 1 implies a decrease of TFP;
- 1985 – 1989: Both productivities decrease, despite new investments in capital and labour. Property (i) of Theorem 2 implies that TFP decreases, irrespective of the value of $\alpha$;
- 1990 – 1993: Less capital and more labour are used, while GDP decreases. Both productivities decrease, with the exception of labour productivity in 1993. In the preceding years of this period TFP thus decreases, irrespective of the value of $\alpha$, which follows directly from expression (24) for TFP;
- 1994 – 1997: Also in this case we find that less capital and more labour are used, but now GDP increases. Both productivities increase, so that TFP increases as well, irrespective of the value of $\alpha$;
- 1998 – 2001: TFP stabilises more or less. The two examples analysed above already showed that TFP increases in 2000, while it decreases in 2001. The development of TFP in 1998 and 1999 depends on $\alpha$.

[85] Statements with regard to an increase or decrease of TFP in time refer to an ordinal scale in this chapter, unless stated otherwise. This also holds for the independence of this behaviour with respect to $\alpha$.

On the basis of the above analysis, we can distinguish two main phases in the development of TFP in Mongolia. Before 1990, new amounts of capital and labour are used in a less productive way in each year, while the reverse occurs after 1993, that is, after the first years of transition from a planned to a market economy.

An important finding of our analysis is that we do not need to know the exact value of $\alpha$ for most inferences (concerning statements on an ordinal scale) about TFP. We therefore obtain stronger conclusions about TFP than, for instance, with Cheng's (2003) regression approach. In addition, the theoretical results about TFP in Section 2 give more insight into the possible causes of its development over the years. On the other hand, the time series contain periods, especially the first half of the 1980s, for which an estimate of $\alpha$ is required in order to draw conclusions about TFP.

A regression approach, such as Cheng has used, makes use of the restrictive assumption that the ratio between the TFP-values in two successive years is constant. This implies that TFP behaves according to an exponential function in time, which is not in accordance with the above analysis of the Mongolian data (notice, for instance, that TFP does not satisfy the property of monotonicity).

Unfortunately, we do not have data about the total amount of wage payments in relation to total GDP in order to estimate $1 - \alpha$. We therefore calculated TFP for different values of $\alpha$, namely 0, ⅓, ½, 0.69 and 1. The value ⅓ is often used for calculating TFP – Aiyar and Dalgaard (2005) concluded that one third is a very good approximation for all countries –, while 0.69 is the value used by Cheng (2003). The development of TFP for the five values of $\alpha$ is shown in

Figure 4-2. Conform Theorem 2 and our conclusions on the period 1980-2001 for Mongolia, the dependence of TFP on $\alpha$ is notable in the period 1980-1984. TFP increases for $\alpha$-values up to ½, while it decreases for the two higher values selected. There is hardly any difference in the behaviour of TFP in the subsequent years as $\alpha$ varies, so that statements about the ordinal behaviour of TFP over these years are robust.

**Figure 4-2. Development of TFP in Mongolia since 1980, based on expression (5), for different values of $\alpha$, and the data in Cheng (2003, Appendix Table 2)**

### 4.3.3 Data quality

***Unreliable data in communist period***

The usefulness of an indicator such as Total Factor Productivity is highly dependent upon the quality and reliability of the data of an economy. With regard to (time series) data, all transition economies face a common problem which is very fundamental: In the communist period, statistics often reflected planned targets rather than the real situation and tended to overstate economic growth as well as the capital stock – both the initial capital stock as well as the ratio of capital stock to GDP – (see, amongst others, Ofer, 1987; Kaser, 2004; Kontorovich, 2001). The underestimation of depreciation inflated the data on capital stock in the former communist countries. Kontorovich demonstrated that for the Soviet Union an adjustment for inflation of the capital stock data leads to an estimated TFP that would have been increasing rather than decreasing in the period from the 1960s to the 1980s.

***Capital stock in Mongolia over-estimated***

If one looks into the situation of Mongolia with the advantage of hindsight, the capital stock (in US$) has indeed been over-estimated due to the fact that until 1990 virtually all investments were done with money borrowed from the USSR. For this and other purposes the government of Mongolia borrowed an amount of 10 billion transferable rubbles. They paid back an amount of 250 million US$ in 2003, which in retrospect suggests that in US$ investments in the years prior to 1990 have been grossly over-estimated. Mongolia – as well as most transition economies – has the additional statistical problem that the methods of gathering data have changed considerably during the transition. Under guidance of international organisations new statistical methods have been introduced to conform to international standards. That means that somewhere in the 1990s there is another break in the time series data.

***Data smoothed and manipulated***

Apart from the usual problems with the aggregation of data, data in Communist countries were smoothed and manipulated while being aggregated regionally or nationally. When using the original data on the firm level, the value of TFP may differ significantly from the manipulated aggregated data. For Mongolia, Total Factor Productivity – both for the overall economy and for the agricultural sector – declines in the decade before the transition. Bayarsaihan and Coelli looked into the original grain and potato production data of 48 state farms in Mongolia in the 14 years prior to the transition and came to the surprising conclusion that, even though the performance of the Mongolian state farms was poor and TFP declined over the whole period, after an initial decline in the 1970s, the farms achieved an impressive TFP growth of 7% per year in the latter half of the period! (See Bayarsaihan and Coelli, 2003). On basis of the official, aggregated data these conclusions could never have been drawn.

***Informal sector excluded from calculation of TFP***

A factor significantly influencing Total Factor Productivity – which is important not only for transition economies but for developing countries as well – is the exclusion of the shadow economy or the informal sector from the data used to calculate TFP. The reason for this is obvious as data from the informal sector are at best unreliable and incomplete, but most of the time they are not available. Thus, the size and productivity of the informal sector have to be estimated. The outcome of these estimations varies considerably with the method used for the estimations. Thus this may lead to serious (measurement) errors in TFP. However, in case of transition economies such as Mongolia, a deindustrialisation and a large-scale redistribution of labour has taken place from the formal sector to the informal sector since the beginning of the transition to a market economy. In Mongolia the informal sector was virtually non-existent in 1990 as the country – unlike most other Communist countries – had not known a "second" economy (Morris, 2001). According to the World Bank, in only 8 years time an estimated 30-35% of Ulaanbaatar's work force was occupied in the informal sector and informal activity was estimated to be in the neighbourhood of 35% of officially recorded GDP (Anderson, 1998)[86]. When the informal economy accounts for roughly one-third of the GDP of a country, it may seriously bias TFP if it is excluded.

***Redistribution of labour from formal to informal sector***

With regard to the measurement of TFP there is on the one hand the question whether capital, entrepreneurial, and technological activities have moved from the formal (state) sector to the informal sector. On the other hand, there is the question of the large-scale redistribution of labour from the formal towards the informal sector which may have various effects upon TFP. When looking into the case of Mongolia, the informal sector tends to be more labour intensive and capital extensive than the formal sector. The informal sector consists of small production units (micro enterprises) in which the advantages of economies of scale cannot be realised. If a Mongolian micro enterprise does grow into a small enterprise – a rare occurrence – and has to take on additional labourers (outside family members), it changes to the formal sector. Thus the movement from the formal to the informal sector is one from large(r) enterprises towards micro enterprises with a very small output. If this trend would be taken into account in the calculations of TFP one would intuitively say that the effect on TFP would show a declining value of TFP as micro-enterprises can never be as efficient as large(r) enterprises as they cannot realise economies of scale – when they can they change to the formal sector in Mongolia – and there is less scope for productivity gains as the activities of micro-enterprises are more labour intensive and have higher labour-to-capital ratios[87].

[86] The enormity of this redistribution is well illustrated by the income data of Mongolian households. Whereas wages and salaries had roughly comprised 85% of average household income in the late 1980s, in 1996 the average household received only 43% of income from wages and salaries (Anderson, 1998, p.13). In the mid-nineties, households in which the head was 'unemployed' received 55% of their monetary income from informal sources, whereas households in which the head was employed received an average of 30% of income from informal sources (Anderson, 1998).

[87] The international standards to define the informal sector exclude households that engage in agricultural activities from the informal sector (Morris, 2001, p.6). This is relevant, as in the agricultural sector of transition countries it is a well-known phenomenon that small plots of land cultivated by families have higher yields per unit of land than the much larger (former) state farms. This is attributable in part to so-called self-exploitation. In the industrial sector, larger enterprises are in general more productive than

Thus a large-scale movement from the formal to the informal sector as has happened in Mongolia should be reflected in a downward trend of TFP as it is a movement away from efficiency.

Mongolia also provides evidence for another effect the shift from the formal to the informal sector may have on TFP. The movement of labourers towards the informal sector in Mongolia is caused on the one hand by reorganisation and closures of enterprises, which leaves people without employment and without significant sources of income as the social safety net in Mongolia is not well developed and Mongolia does not have the resources to support all the unemployed out of the public purse. On the other hand, however, due to the incentive structures in the formal and informal sector, there is also a large voluntary movement to the informal sector as it pays much better than many jobs in the formal sector (especially as compared to government-paid jobs such as in the health, education and judicial sector). Moonlighting – having a job in the formal sector, but working in the informal sector as well – is a common phenomenon in Mongolia, but there are also many highly skilled people who give up their jobs as doctors, nurses and teachers in the formal sector to earn a much better wage as taxi-driver, trader or cleaner in the informal sector. The Mongolian informal sector is characterised by the fact that people tend to be highly educated (Morris, 2001); people are much more skilled than the level of their micro-enterprise requires.

***Influence of movement to informal sector on TFP***

Theoretically seen this voluntary movement of workers towards the informal sector could potentially influence TFP negatively in three ways and positively in one way, that is, if the informal sector would not be excluded from the calculation of TFP. In the first place, it would influence TFP negatively because people change jobs from more efficient enterprises to less efficient enterprises (because of the change from large(r) enterprises to micro-enterprises). In the second place, the voluntary movement towards the informal sector concerns in general only the most productive workers, those with initiative and with enough skills. Others will not voluntarily leave the formal sector for a better paid, but insecure job in the informal sector. This implies a loss of know-how, skills and efficiency in the formal sector. This phenomenon should also influence TFP in the long run as for instance doctors who have left the health sector for ten or fifteen years cannot return to the health sector anymore as in that time they may have lost their skills and their knowledge will be outdated. Thus this trend is irreversible unless people are re-educated. In the third place, some sectors in Mongolia, such as the health sector, have difficulty to maintain their productivity and suffer a loss of efficiency as too many (skilled) workers have left the sector, which ought to have a negative effect on TFP. In the long run the loss of efficiency in the health and education sector may actually show up in the data as labour productivity and TFP will go down[88] (unless, statistically seen, all those who suffer from the decline of efficiency in those sectors are active only in the informal

smaller enterprises due to economies of scale. Micro-enterprises tend to invest less in capital (goods) and more in labour.

[88] Whereas once education and health services were universal, at this moment it is predominantly the non-poor part of the Mongolian population that continues to be served by these sectors despite government interventions to keep the services accessible for the poor as well.

sector. It would still imply a loss to the economy as a whole, but it would not be measured in the official TFP). Due to the peculiarities of income accounting, the last two effects will not be reflected in TFP. Many structural system changes such as an increase or decrease in barter trade are not measured in national income data. The output of public services is measured on the basis of the wages of public servants, and education and health are measured on the basis of inputs. Formally, the higher wages paid in the informal sector imply a higher efficiency, and thus a higher value of TFP. Thus, in the formal sense, as the taxi-driver is rewarded with a higher wage, it is more efficient to have a doctor working as a taxi-driver or cleaner in the informal sector than as a doctor in the formal sector. However, output in the market is measured in (free) market prices, whereas government wages were and are still strictly controlled, which leads to an imbalance in the measurement of (the efficiency of) the governmental sectors as compared to the other (market) sectors. The conclusions drawn on the basis of this measurement imbalance are therefore not necessarily valid. And as we have seen, the efficiency of the overall economy does suffer from the shift of highly skilled educational, health and other workers to the informal sector but this is not measured in TFP.

***Reductions in TFP could be a longer-term phenomenon if measured accurately***

The fall in output in many transition countries in the initial years of their transition to a market economy in combination with the phenomena with regard to the informal sector described above provide evidence that a structural loss of knowledge, skills and efficiency can occur in a country. As the described phenomena largely take place outside the formal sector, they are usually not reflected in the value of Total Factor Productivity. The described phenomena also shed another light on the assumption of Griliches and Lichtenberg (1984) – introduced to deal with some data problems – *that "true" productivity can only improve, so that measured reductions in TFP only reflect short-term fluctuations*. Unless a time-span of fifteen years, and more, is regarded as a short term, productivity can decline structurally and the phenomena accompanying the redistribution of labour from the formal economy towards the informal economy provide evidence that reductions in TFP could in fact be a longer-term phenomenon if it were measured accurately.

***Barter trade and the value of TFP***

A specific factor that influences the value of TFP for Mongolia is the (large-scale) existence of barter trade, especially in the rural areas. When calculating TFP, illegal and criminal activities are usually excluded from the calculation. As barter trade is assumed to be undertaken as a strategy of tax evasion, it falls under this category. In Mongolia barter trade in general is not undertaken as a means to evade taxes, however, but it is rather a direct consequence of macroeconomic austerity measures, that restricted the money supply and thus caused severe cash-flow problems. This gave rise to demonetisation of exchange and caused a significant increase in barter trade. The outcome was serious efficiency losses, in terms of flexibility and price competition (Odgaard, 1996). As barter trade in rural areas is standard in Mongolia, it is a serious omission that it cannot be taken into account in the calculation of TFP. It means that a significant part of the economy remains outside the calculations of TFP. This problem cannot be solved as there are no reliable estimations of the amount of barter trade in

Mongolia. It is furthermore difficult to predict whether this omission would have influenced TFP positively or negatively, though one would tend to expect that it would lead to an over-estimation of TFP as all larger enterprises – which can profit from economies of scale and are thus more efficient – are situated in urban areas[89].

***Problems with the concept of TFP for communist economies***

When calculating Total Factor Productivity for the formerly communist countries, there is furthermore a problem of a theoretical nature. The concept of TFP is based on the assumptions of perfect competition and marginal productivity pricing. The communist economies were greatly distorted with regard to these assumptions. Prices of both capital and consumer goods and wages did not – and sometimes still do not – reflect "market or marginal prices". The values for TFP have always been corrected on the basis of various assumptions with regard to the available data and there has always been discussion on the exact methods to be used to correct the data. Generally, one can say that the values of TFP can be compared among the communist countries of Central Asia for example, as it can reasonably be assumed that these countries roughly had the same economic structure and used the same accounting techniques and thus that the distortions will work in the same direction. It is, however, more difficult to compare the TFP-values "over the various systems" before and after the transition in 1990. It also means that there are difficulties in comparing the TFP values for the ex-communist countries before, during and after the transition. In most transition countries prices have been liberalised in the first decade of transition[90]. During the period of the liberalizations one can say that the prices have moved in the direction of marginal productivity pricing. However, in Mongolia it took more or less from 1990 until 1996 before all prices were liberalised and still some prices are being perceived as being subsidised by the international organisations as they are still much lower than in neighbouring countries. This movement of prices definitely has an influence on TFP and it moreover has a different effect in each subsequent year in Mongolia as every year more and more prices were moving towards market prices. It is however difficult to say how it ought to have influenced TFP.

### 4.3.4 Sensitivity analysis

The conclusions about TFP that were drawn in Section 4.3 are based on the Cobb-Douglas function for GDP. These conclusions were found to be robust with regard to the parameter $\alpha$: the ordinal behaviour of TFP in time shows a small dependence on $\alpha$ for the economic situation in Mongolia. This finding strengthens the conclusions about TFP considerably in comparison with Cheng's study (2003). However, the conclusions about TFP are based on the available data and on the value of 0.06 for the depreciation rate $\delta$. As Cheng (2003) did not give precise arguments about his choice for the depreciation rate and since the data may be distorted, as we saw in Section 4.3.3, we will perform a sensitivity analysis in order to quantify the effect of different sources of measurement error in the data and of uncertainty in the value of the depreciation rate.

[89] A fact that also underlies this expectation is that in rural areas the herders with large herds do tend to get paid in cash by traders when they sell animals, whereas the herders with smaller herds are paid in kind. Thus the larger, more productive "agricultural units" are accounted for in TFP.

[90] Uzbekistan is a notable exception.

The most important aspect of this analysis is that we extend the scope of the sensitivity analysis to cover all parameters and variables of the Cobb-Douglas function. Most studies focus only on variations in the estimated parameter values and do not involve possible measurement errors in the data (GDP, capital and labour). We carry out a sensitivity analysis[91] by setting parameters and data variables at certain values, for which we quantify their individual and combined effect on TFP.

We first formalise the sensitivity analysis by transforming the components of the Cobb-Douglas function in order to account for the possible sources of error and uncertainty, and their size. Next, we give the results of the sensitivity analysis for specific combinations of variations around the original values on GDP, capital and the depreciation rate.

As sources of variation we introduce the nonnegative parameters $c_Y$, $c_K$, $c_\delta$, $c_L$ and $c_q$, which we use to vary GDP, capital, the depreciation rate, labour and the human capital index, respectively. The sensitivity analysis for the Cobb-Douglas function thus will have the following components:

- The depreciation rate becomes $c_\delta\delta$, where $\delta$ is fixed at the original value 0.06;
- We transform GDP $Y_t$ into $c_Y^{\tau-t+1}Y_t$, where $\tau$ denotes the most recent year of the data record (i.e., $\tau = 2001$ in this study);
- We apply the same transformation to investments $I_t$ in year $t$, yielding $c_K^{\tau-t+1}I_t$;
- For labour we assume the same error for each year, so that labour in year $t$ becomes $c_L L_t$;
- For education we assume the same error for both values on years of schooling in 1989 and 1998; we apply an error of the size $c_q$ to both data values.

Substitution of the above transformations for the depreciation rate and investments in (22) yields the following expression for capital $K_t^*$ in year $t$:

$$K_t^* = \sum_{i=0}^{t} (1 - c_\delta \delta)^{t-i} c_K^{\tau - i + 1} I_i, \tag{25}$$

which can be written in the more convenient form

$$K_t^* = c_K^{\tau - t + 1} \sum_{i=0}^{t} \{c_K (1 - c_\delta \delta)\}^{t-i} I_i. \tag{26}$$

The transformations of the parameters and data variables are formalised in such a way that we obtain the original values when the sensitivity analysis parameters are set equal to 1. The original data and parameter values underestimate the true values if the sensitivity analysis parameters are greater than 1, whereas the true values are overestimated when

[91] We perform a sensitivity analysis instead of an uncertainty analysis, because the latter requires probability distributions of parameter and data values. As we do not have accurate information about the shape of such distributions, we carry out a sensitivity analysis in the conventional sense.

these parameters are smaller than 1. For instance, $c_K < 1$ applies to situations where the official data overestimate capital stock, which is an interesting case in the light of the situation in Mongolia (Section 3.3).

The effect of variations of the original data and parameter values on TFP in year $t$ can be quantified by substituting the transformed quantities in the expression

$$A_t = \left(\frac{Y_t}{K_t}\right)^{\alpha} \left(\frac{Y_t}{q_t L_t}\right)^{1-\alpha},$$

which gives TFP $A_t^*$:

$$A_t^* = \frac{c_Y^{\,\tau-t+1} Y_t}{c_K^{\,(\tau-t+1)\alpha} \left[\sum_{i=0}^{t} \{c_K(1-c_\delta\delta)\}^{t-i} I_i\right]^{\alpha} \left[c_q q_t c_L L_t\right]^{1-\alpha}}. \qquad (27)$$

Although variations in labour and human capital affect TFP, they do not affect conclusions about comparisons between TFP-values in successive years, because we assumed constant relative error sizes over time for both variables[92]. We will therefore ignore variations in labour and human capital, and we set $c_L = c_q = 1$. Variations in the other three sources of (27) have the following effects on TFP in any year $t$:

- If $c_Y$ increases (decreases), then TFP increases (decreases) as well;
- If $c_K$ increases (decreases), then TFP decreases (increases);
- If $c_\delta$ increases (decreases), then TFP also increases (decreases).

The goal of this sensitivity analysis is to investigate to which extent the conclusions about TFP, as presented in Section 4.3 for the original data, change under individual and simultaneous variations around the original data and parameter values. For instance, if TFP was shown to grow in year $t + 1$ with respect to year $t$, so that $A_{t+1}/A_t > 1$, does this still hold under data and parameter variations? Before proceeding with the implementation of the sensitivity analysis and the results, we make some analytical considerations that will be helpful in interpreting the final results.

In formal terms, the TFP-ratio will be transformed in the sensitivity analysis into the ratio

$$\frac{A_{t+1}^*}{A_t^*} = \frac{c_K^{\,\alpha}}{c_Y} \left[\frac{\sum_{i=0}^{t} \{c_K(1-c_\delta\delta)\}^{t-i} I_i}{\sum_{i=0}^{t+1} \{c_K(1-c_\delta\delta)\}^{t+1-i} I_i}\right]^{\alpha} \left(\frac{K_{t+1}}{K_t}\right)^{\alpha} \frac{A_{t+1}}{A_t}, \qquad (28)$$

which follows from (27). The coefficient of $A_{t+1}/A_t$ in (28) determines the size of the change in relative growth or decline of TFP.

[92] We have assumed constant relative error sizes over time for labour and human capital as labour and education are relatively easy to measure and thus are – at least in case of Mongolia – less likely to have measurement errors that vary significantly from year to year.

In order to interpret the results presented later in this section, let us consider (28) for the situations $t = 0$ and for increasing $t$. As $t \to \infty$, the ratios within brackets in (28) go to 1, since the sums are convergent[93]. This implies that (28) behaves as $(c_K^{\alpha}/c_Y)(A_{t+1}/A_t)$ as $t$ increases. This limit behaviour is independent of $c_\delta$ and $\delta$, as the effects of capital depreciation decrease with time. This analysis implies that GDP becomes the dominant factor as $t$ increases. Variations in GDP have a linear effect on the relative growth or decline of TFP.

For $t = 0$, expression (28) is equal to

$$\frac{A_1^*}{A_0^*} = \frac{1}{c_Y}\left\{\frac{(1-\delta)I_0 + I_1}{(1-c_\delta\delta)I_0 + \frac{I_1}{c_K}}\right\}^{\alpha}\frac{A_1}{A_0}. \tag{29}$$

Also in this case we observe that GDP-variations can have a major impact on the conclusions about TFP. The parameter $c_K$ will have small effects for small $t$ (Capital stock $I_0$ in 1980 is much greater than investment $I_1$ in 1981, and will thus largely determine (29).) Variations in investments will have a larger effect when $t$ increases, as we saw in the preceding analysis. Variations in the depreciation rate will have a larger effect than variations in investments for small $t$. If we set $c_Y = c_K = 1$, then (29) becomes

$$\frac{A_1^*}{A_0^*} = \left\{\frac{(1-\delta)I_0 + I_1}{(1-c_\delta\delta)I_0 + I_1}\right\}^{\alpha}\frac{A_1}{A_0}, \tag{30}$$

which is equal to about $1.058^{\alpha}A_1/A_0$ for $c_\delta = 2$. Note that the effect of variations in the depreciation rate on the TFP-ratio $A_1/A_0$ becomes larger when $\alpha$ increases, as it is the income share of capital.

Finally, we note that the sensitivity analysis formalised in this section includes a special case that is worth emphasising. A sensitivity analysis could also be performed by applying equal yearly variations for every variable and parameter, that is, $c_Y Y_t$ and $c_K I_t$ instead of $c_Y^{\tau-t+1}Y_t$ and $c_K^{\tau-t+1}I_t$, for every year $t$. The effects of such variations on comparisons of TFP between successive years can be quantified by setting $c_Y = c_K = 1$ in (28), since variations in both GDP and investments cancel out when making comparisons between different years. The results of such an analysis will therefore only depend on variations in the depreciation rate. From the previous analysis, it thus follows that conclusions on TFP-growth will remain unchanged as $t$ increases, while for $t = 0$ we obtain (30).

We used the following variations around the original parameter and data values: $c_\delta = ½$ and 2, $c_K$, $c_Y = 0.975$ and 1.025. The lowest and highest values of the depreciation rate are therefore 0.03 and 0.12, which cover the range 0.04-0.1 used in other studies. The

[93] Investments have an upper bound and $c_K(1 - c_\delta\delta)$ and $1 - \delta$ are smaller than 1 and nonnegative. Each of the four series for capital stock therefore has an upper bound that converges, so that the original series are also convergent.

variations in GDP and investments become larger when going back in time. In this way we can account for larger errors in data acquisition for remote years, due to less sophisticated statistical methods and other sources mentioned in Section 4.3.3. A value of $c_Y = 1.025$ implies that the variation around GDP in the reference year $\tau = 2001$ is 2.5%, while for 1980 the variation is about 72%. We believe that this range is sufficiently large for investigating the sensitivity of the conclusions about TFP.

We will first quantify the effect of each individual parameter variation by fixing the other sensitivity parameters at the value 1 (i.e., no change in the original values of the other parameters and data variables of the Cobb-Douglas function). Next, we will quantify the largest variations in TFP by combining the largest individual variations in all parameter and data values. At the end of this section, we will also say something about possible effects on TFP when (a part of) the informal sector is included.

The results for the individual variations in GDP, investments and depreciation rate are shown in Figure 4-3. One parameter or variable at a time is varied, while the other factors are fixed at their original values. The analytical properties of the behaviour of the TFP-ratio (28) can be easily recognised in Figure 4-3: variations in the depreciation rate become less important as time passes, the effects of variations in investments and capital stock are negligible for the most remote years and are very small in general, while variations in GDP have the same effect on the TFP-ratio for every year. This shows that the underestimation of the depreciation rate and the overestimation of the capital stock-to-GDP ratio, as described in Section 4.3.3, hardly have any effect on the original conclusions about TFP.

The results also confirm that variations in GDP have the largest effect on the original conclusions regarding TFP-behaviour. The line $A_{t+1}/A_t = 1$ in Figure 4-3 (a)-(c) denotes the situation where TFP does not change in time. A value greater than 1 means that TFP increases, while a value smaller than 1 means that TFP decreases in the next year. The variations in the two directions around the original ('baseline') values for GDP show that the original conclusions about the TFP-ratio $A_{t+1}/A_t$ change six times (i.e., in six years) in sign for both variations. This happens only twice for the two depreciation rate variations, while variations in the original investment values hardly have any effect.

The years in which conclusions are changed coincide for the three factors: when an arbitrary parameter is set at its largest variation, conclusions about the sign of TFP will change in 6 out of 21 years at most. For the years 1990 -1992 and 1994 -1997, the original conclusions do not change at all.

This picture does not change much for other values of $\alpha$. The results become more robust when $\alpha$ decreases, as the effects of GDP variations on the TFP-ratio are independent of $\alpha$ (see (28)), while the influence of depreciation rate and capital stock become smaller. When $\alpha$ increases, variations in investments still show negligible effects. The same holds for variations in the depreciation rate with regard to the TFP-ratios in recent years, which is in agreement with the above theoretical analysis. There is a larger effect, however, for the period 1980-1985 when the depreciation rate is changed from 0.06 to 0.12. TFP decreases in each of these years for the baseline value of 0.06 when $\alpha$ = 0.69 (and of course also for greater values of $\alpha$), but increases in each year when the depreciation rate is set at 0.12. Of course, the conclusions are unchanged when setting the depreciation rate at 0.03, since capital stock increases, while GDP remains the same.

The results of the last paragraph invite us to consider the effects of combined variations in the three factors. Figure 4-4 shows the largest combined effects in both directions around the original values for three values of $\alpha$. The results are comparable to the results for the individual variations Figure 4-3); the conclusions about TFP-change in successive years change between five and eight times with respect to the baseline values. A notable exception is the situation where combined variations in GDP, investments and the depreciation rate give rise to the largest increase of the TFP-ratios in successive years when $\alpha$ = 0.69 (Figure 4-4). The original conclusions about TFP-change then switch in sign in 12 out of 21 years of the time series.

Overall, the conclusions about TFP-development as stated in Section 4.3.2 are robust under individual variations in depreciation rate and $\alpha$ in particular but are less so for GDP. When individual variations are combined unfavourably, the sign of TFP-change changes quite often and the magnitude of the change may be considerable. Even in this case, however, the original results are very robust for the years 1990-1992 and 1994-1997, for which we can reliably stick to the conclusions of Section 4.3.2 (i.e., TFP decreases during 1990-1992, while TFP increases during 1994-1997). The results of the sensitivity analysis furthermore show that GDP has the largest effect upon the original conclusions. This analysis thus shows that an accurate assessment of especially GDP will considerably reduce the uncertainty in TFP-development, while accurate estimates of the depreciation rate, along with $\alpha$, are important for remote years.

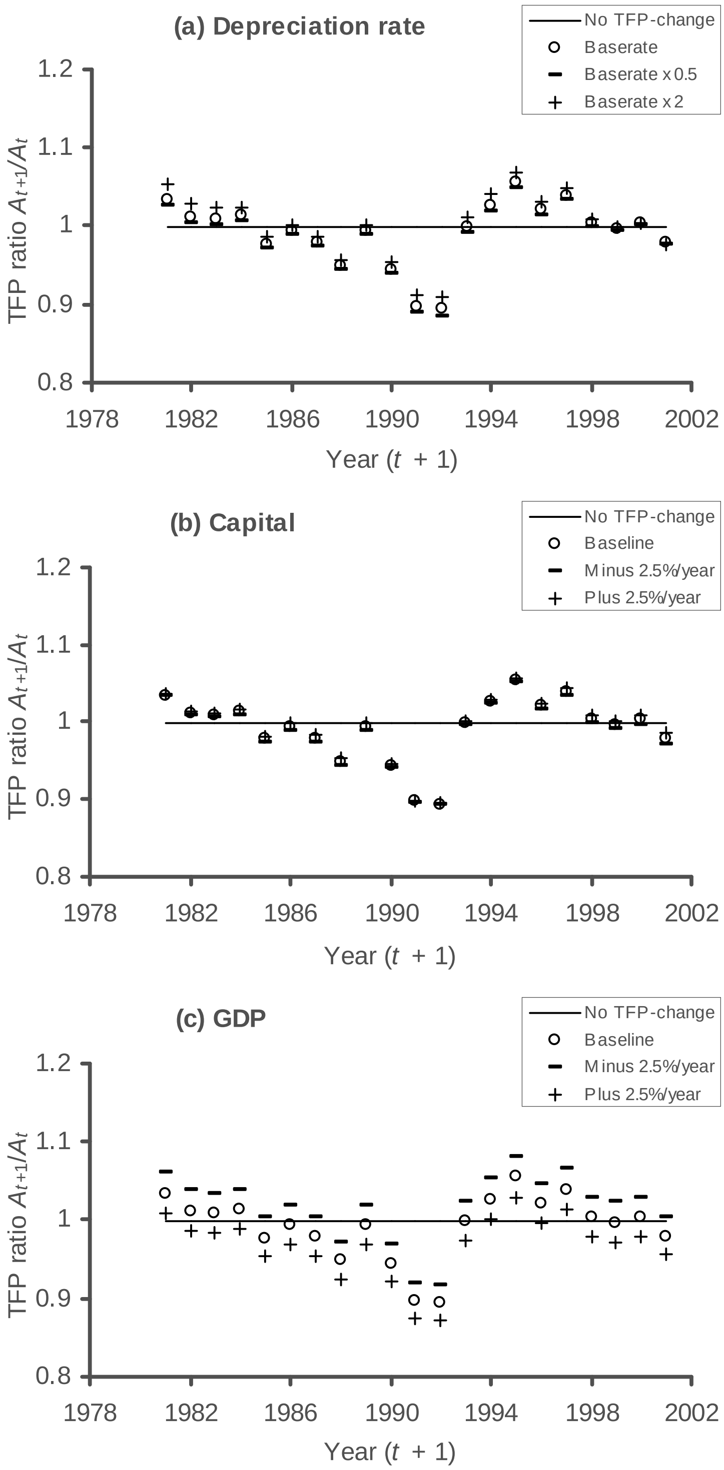


**Figure 4-3. Effects of individual variations in GDP, investments and depreciation rate on the ratio of the TFP-values for successive years, with $\alpha$ = ⅓. 'Base rate' and 'baseline' refer to the original values.**

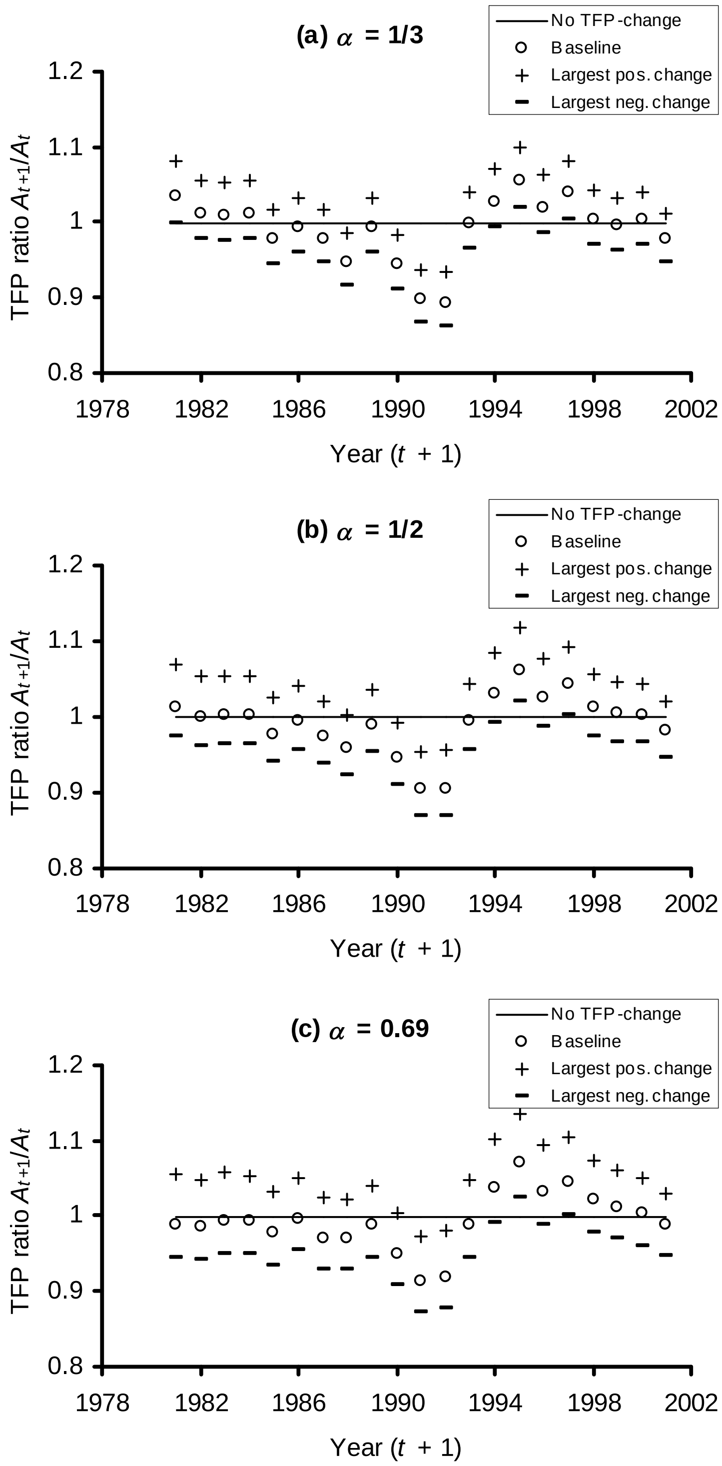


**Figure 4-4. Effects of combined variations in GDP, investments and depreciation rate on the ratio of the TFP-values for successive years, for $\alpha$ = ⅓, ½ and 0.69. The largest deviations in both directions from the original (baseline) values are shown: $c_Y$ = 0.975, $c_K$ = 1.025, $c_\delta$ = 2 (+), and $c_Y$ = 1.025, $c_K$= 0.975, $c_\delta$ = ½ (-).**

Finally, we also showed that in periods of structural changes, such as the growth of the informal sector and the increase in barter trade in the period 1990-1995 in Mongolia, TFP may significantly deviate from the measured values. Once the economic situation has stabilised and phenomena as barter trade and the informal sector do not significantly change in magnitude anymore from one year to the other, their influence in TFP-change cancels out.

## *4.4 TFP in a simulated economy*

One of the goals of this chapter is to find a satisfying explanation for the increase of TFP in Mongolia after the transition to a market economy. The analyses in the previous section showed that the increase of TFP during the years 1994 –1997 is very robust in the sense that it is independent of $\alpha$ (Section 4.3.2) and that this finding is hardly affected by the largest combined variations in GDP, depreciation rate and investments (Section 4.3.4). To which factors can this behaviour be attributed? Are the changes in TFP attributable to increases in efficiency due to for example privatization and liberalization, or are they caused by other factors such as the closing down of enterprises or the excess of labour due to retention of labour during the previous period of recession, thus enabling enterprises to increase production without the necessity of investing in capital and labour.

In the wake of the Mongolian privatizations in the early nineties, many enterprises closed down. These were not necessarily the least efficient enterprises nor the least promising. The largest enterprises have basically not been closed down. An important reason for that was that the largest enterprises were the main generators of foreign exchange and thus added to the much-needed foreign reserves of Mongolia. Moreover, the large enterprises had the important social task of employing many Mongolians. Closure of (some of) the largest enterprises in the nineties would have amounted to a social catastrophe.

In order to investigate what factors may have contributed to the behaviour of Mongolian TFP in the nineties, we have set up a number of simulations to see what happens with TFP when, in a closed economy, a certain percentage of enterprises closes down. The analysis of the mathematical model in Section 4.2 describes some consequences for TFP in specific situations (Theorems 2 and 3). The simulations in this section permit us to investigate this hypothesis more thoroughly and allow us to look at an imperfect, but more realistic situation in which not all of the usual assumptions with regard to the model underlying TFP need to be fulfilled. Thus marginal pricing is not a necessity in a simulation, neither is it a necessity that only the least efficient enterprises close down. Especially the latter characteristic is important as the closure of enterprises in Mongolia appears to be compatible with a random way of closing enterprises –with the exception of the largest enterprises that did not close down- rather than with closing enterprises based on the efficiency of those enterprises.

In our simulation we considered a closed economy with $n$ = 1000 enterprises. They are represented by triplets ($Y_1$, $K_1$, $L_1$), …, ($Y_n$, $K_n$, $L_n$), each with production $Y_i$, capital stock $K_i$ and labour $L_i$, which satisfy the Cobb-Douglas production function. The enterprises were subdivided into three classes: ‘large’, ‘medium’ and ‘small’.

Production, capital and labour of each enterprise were simulated independently of each other from uniform distributions, with values within the ranges [0, 10 000], [0, 2000] and [0, 500], respectively. Those values can be rescaled according to certain scaling constants in order to represent realistic values. This will, however, not affect any of the conclusions that will be drawn about comparisons between old and new TFP-values, as such comparisons are invariant under all ratio-scale transformations of production, capital and labour in the Cobb-Douglas production function (Theorem 1). The TFP of the economy, that is, for the ensemble of *n* enterprises, follows from the Cobb-Douglas function as the quantity

$$A = \frac{Y}{K^{\alpha} L^{1-\alpha}}, \tag{31}$$

where *Y*, *K* and *L* are the sums of production, capital and labour, respectively, over the *n* enterprises.

To investigate the effect of closures of enterprises on TFP we have simulated different scenarios. In scenarios (i) to (iii) we randomly closed a certain percentage of enterprises in order to simulate an economy that resembled the Mongolian economy in the first decade of the transition. In scenarios (i) and (ii) we included the observation of the Mongolian situation that the largest enterprises were kept open. To this end, we simulated scenarios (i) to (iii) in a probabilistic model, in which enterprises have a probability *p* of closing down. This probability was varied between 0.1 and 0.9, with intermediate values chosen with increments of 0.1. Enterprises stayed open or were closed independently of other enterprises. Scenario (iv) investigates the effect on TFP when a certain percentage of the least efficient enterprises is closed, that is, the enterprises with the smallest TFP-value. The scenarios are as follows:

(i) The 20 enterprises with the largest production were not closed. Enterprises belonging to the other 98% were closed with probability *p*, which was the same for all these enterprises;
(ii) The 10 enterprises with the largest production were not closed. Enterprises belonging to the other 99% were closed with probability *p*, which was the same for all these enterprises;
(iii) Each of the 1000 enterprises had the same probability *p* of being closed;
(iv) Of the 1000 enterprises, the least efficient enterprises were closed first.

Scenarios (iii) and (iv) served as benchmarks for scenarios (i) and (ii). Scenario (iii) enabled us to answer the question whether an increase in TFP is more likely to occur when only small and medium enterprises are considered for closure and the largest enterprises remain open, rather than considering all enterprises. Scenario (iv) gave us the opportunity to compare the results of our simulations with the ideal type outcome of TFP-theory.

For each of the nine closing probabilities *p* = 0.1, 0.2, …, 0.9, we generated 1000 samples. An individual sample was generated by simulating, for every enterprise,

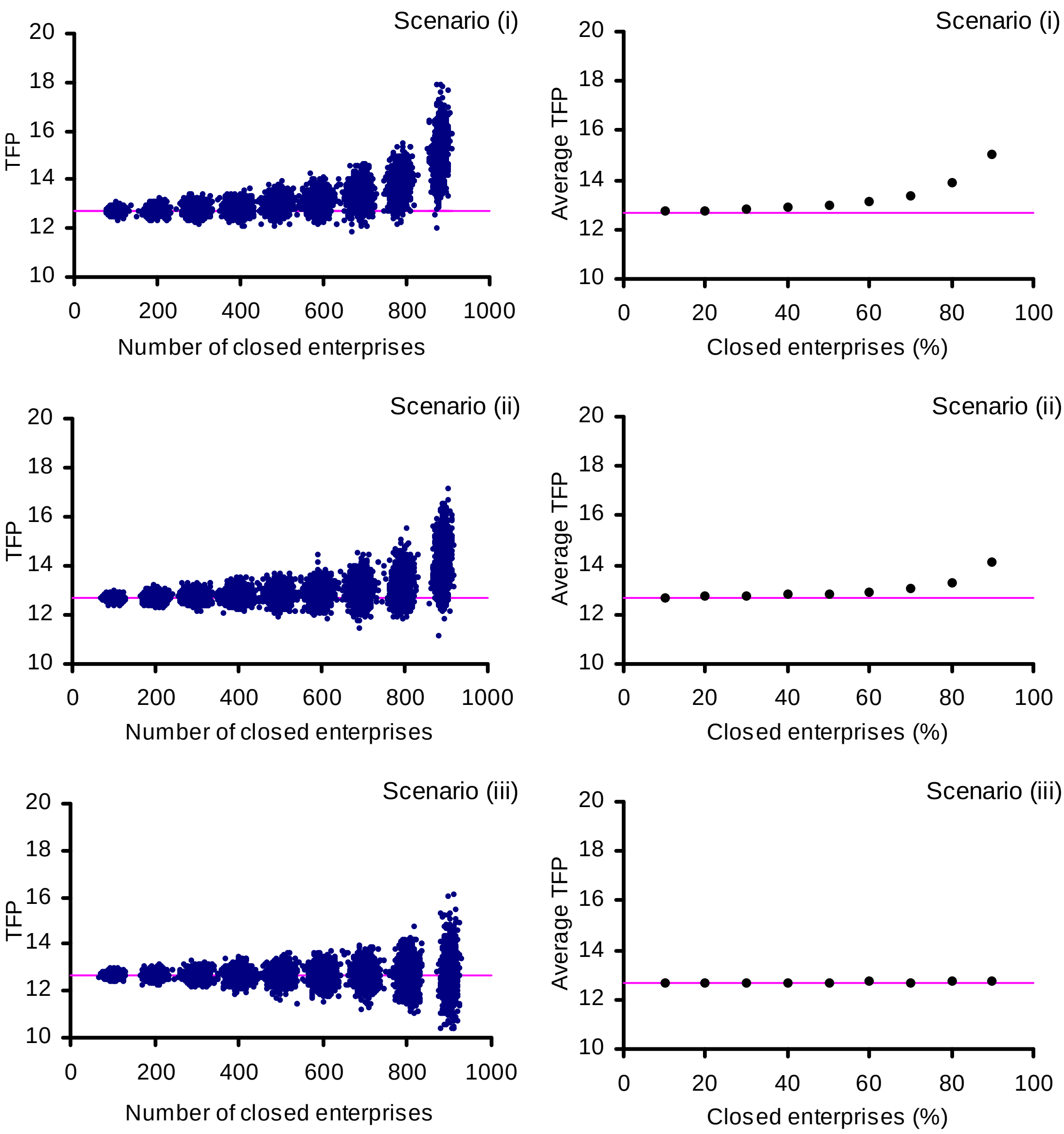


**Figure 4-5. Simulation results for scenarios (i)-(iii). The leftmost figures contain the TFP-values of each individual sample. The nine clusters each contain 1000 samples that correspond with the nine closing probabilities, ranging from p = 0.1 (leftmost cluster) to p = 0.9 (rightmost cluster). The rightmost figures contain the average TFP-values over 1000 samples for each of the nine closing probabilities. The horizontal lines denote the old TFP-value of the economy before the simulated reforms.**

whether it was closed or not according to $p$. Of course, in scenarios (i) and (ii), the 2- and 1-percent largest enterprises were not closed. For each of the 1000 samples, the TFP-value of the simulated, transformed economy was calculated according to (31) for the enterprises that remained open.

The results of the simulations of the three scenarios are shown in Figure 4-5, for which we set $\alpha = ⅓$.[94] The results can be compared with the old TFP-value, which is indicated by the horizontal line in each plot. The results show that: (1) scenario (iii) does not lead to an increase of TFP on average as the number of closed enterprises increases, (2) TFP increases on average as the number of closed enterprises increases, when a certain percentage of the largest enterprises, that is, enterprises with the largest production, are not closed (scenarios (i) and (ii)), (3) the probability that the new TFP-value will be greater than the old value increases and goes to 1 as the closing probability increases towards 1 under scenarios (i) and (ii).

Results (2) and (3) remind of property (i)-(a) of Theorem 3 in Section 2, which implies that for a subset of enterprises that is more capital-productive and labour-productive than the remaining set of enterprises, the TFP of the first subset exceeds the TFP of the overall economy that includes all enterprises. In other words, when a certain amount of enterprises are closed such that the remaining enterprises have a greater capital- and labour-productivity on aggregate than was the case with the former economic situation, then TFP will increase. This situation is comparable to the simulations under scenarios (i) and (ii).

The yearly increase of TFP in Mongolia in the years 1994-1997 ranges from 2% to 5.5%, for $\alpha = ⅓$. The situation in Mongolia in this period shows that GDP increases, while capital decreases slightly and labour stays more or less at the same level. When this would not be a business cycle fluctuation (i.e. retention of labour during a period of recession), this might point at a small number of middle-sized enterprises being closed, possibly in combination with a fairly large number of very small enterprises. The part of the simulations that we will therefore use for making comparisons with the data of Mongolia is the subset of samples generated for the smaller values of the probability $p$ of enterprises closed.

The simulation results show that a fully randomised closing policy as described by scenario (iii) does not coincide with the actual economic development in Mongolia described above. Neither does scenario (iv), however, in which only the least efficient enterprises are closed. In scenario (iv) TFP rises quite impressively with the closure of more and more enterprises, but at the same time GDP decreases in concordance as is illustrated in Figure 4-6. Scenario (i) gives a better match with the data, although it still underestimates the actual four-year increase of 11.9% considerably. The simulation results show that TFP increases on average with about 1.6% in this scenario when 40% of the enterprises are closed, with TFP-increases of up to 8% emerging from individual samples. In Figure 4-6 the behaviour of TFP as a function of the percentage of open

[94] For the results and conclusions of our simulations, the exact value of $\alpha$ does not make a difference. We set $\alpha = ⅓$ for our simulations based on literature such as that of S. Aiyar and C. Dalgaard (2005).

enterprises, with increments of 10%, is presented. When comparing scenario (iv) with scenarios (i) and (ii) it is evident that TFP increases considerably and in a very rapid fashion under scenario (iv). When 10% of the enterprises are closed, TFP has already increased with 12.2%. When 20% is closed, TFP shows an increase of 24.7%, which becomes 54.1% when 40% of the enterprises are closed. Such drastic changes are not observed in the data. The results in this section suggest a scenario for Mongolia in which a small percentage of the enterprises with the largest production were not closed, which is possibly greater than the 2% in scenario (i), while among the remaining enterprises both non-viable and viable enterprises were closed down in a random fashion.

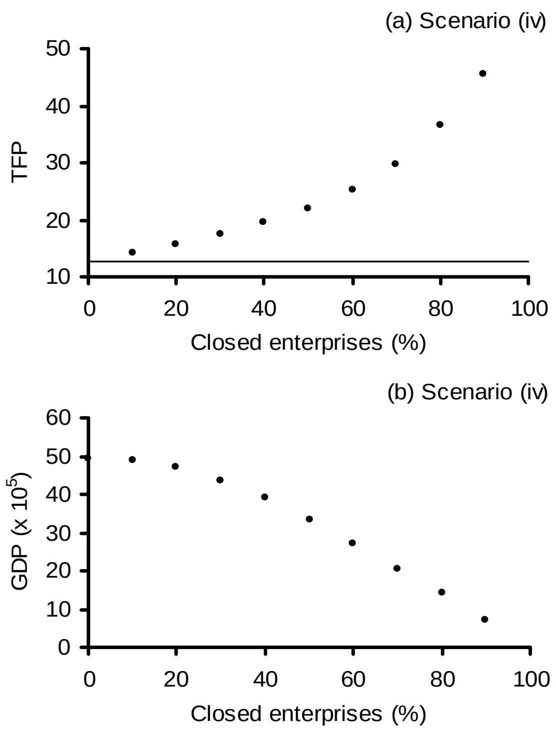


**Figure 4-6. (a) TFP as a function of the percentage of closed enterprises in scenario (iv), in which enterprises are closed on the basis of smallest TFP. The horizontal line denotes aggregate TFP for the 1000 enterprises; (b) Corresponding behaviour of GDP.**

### 4.4.1 Sensitivity analysis

As in this simulation production, capital and labour of each enterprise were simulated independently of each other from a uniform distribution, a sensitivity analysis was performed to see whether the results were robust when a different distribution (β-distribution) was used or when the sums of production, capital and labour over all enterprises in each size class (i.e. large, medium and small) were restricted. Different scenarios were used with regard to the size classes. Some scenario’s that reflect the communist situation at the beginning of the nineties (many large and middle sized enterprises) were used as well as the Western European scenario[95]. Though the Western European scenario was not realistic for the communist countries at the beginning of the transition in 1990, it was used as it is certainly possible that the situation of Mongolia and other transition economies will resemble that scenario in the near future. With the exception of the Western European scenario, the results of the simulations proved to be robust.

## *4.5 Conclusions*

This chapter was inspired by the analysis of the Mongolian economy by Cheng (2003) using the Cobb-Douglas production function. It looked into several mathematical properties of the Cobb-Douglas function and into the consequences of these properties for Total Factor Productivity. The outcomes were used to gain a better insight into the development of efficiency and TFP for the transition economy of Mongolia. The most interesting (and obvious) conclusion with regard to the properties of the Cobb-Douglas function and TFP is that in the commonly occurring cases where both capital- and labour-productivity increase or decrease, the movement of TFP is independent of the value of $\alpha$. Not only for the transition economy of Mongolia but for most economies this is an important finding as often capital productivity and labour productivity move in tandem. It means that in all those cases, in which TFP-growth in time is considered on an ordinal scale, we do not need to bother about estimating alpha and can thus avoid a time consuming calculation. Hence, in these cases the estimation of $\alpha$ by Cheng (2003) was unnecessary.

Another interesting conclusion is the corollary of Theorem 2, property (iii), that an unbalanced and less uniform distribution of capital and labour over the enterprises in an economy affects TFP negatively. From a managerial point of view this is an interesting conclusion as it makes decisions with regard to the inputs easier as even without knowledge of GDP and $\alpha$ one can increase efficiency in an enterprise by balancing the inputs of capital and labour.

The quality of the data may have a significant impact on the sign and value of TFP-change, as well as structural changes in an economy, such as the growth of an informal sector or a change in the amount of barter trade. Once the economic situation has stabilised and the magnitude of the informal sector and barter trade do not change significantly anymore, the exclusion of those phenomena from the calculation of TFP does not really influence the relative change in the value of TFP anymore. In transition economies and developing countries both phenomena mentioned above play an important role in the economy and due to the dynamics of these economies, where these phenomena may change considerably from one year to the other, this may lead to quite a large understatement or overstatement of TFP. Structural measurement errors of more or less

[95] The communist scenario, that in our opinion approached the real situation in Mongolia, was 5% large enterprises, 15% medium enterprises and 80% small enterprises. Because of the lack of precise data, other “communist” scenario’s with many large and medium sized enterprises were used as well, which basically yielded the same results. The Western European scenario was 1% large enterprises, 9% medium enterprises and 90% small enterprises.

the same sign and magnitude every year cancel out in the calculation and comparison of TFP-values for different years and are thus unimportant. The value of $\alpha$ was not important for statements about TFP development in most years and deviations in the depreciation rate and in investments were also found to be less important for the value and direction of TFP-change. Thus, in a period when the economy – and especially the informal sector – changes structurally or when the quality of data is doubtful or changes considerably, Total Factor Productivity should not be used as an indicator of efficiency as basically in such a period its value is meaningless.

An accurate measurement of GDP is of crucial importance to the use of TFP as a measure of efficiency as it is the most influential single factor. Therefore, in economies that are in the process of structural change, be it the growth or decline of an informal sector, of barter trade, or of any other structural change, TFP as a measure of efficiency should be used with caution. Furthermore, for economies with unreliable statistics in combination with the peculiarities of income accounting, the present use of TFP may lead to a distorted picture of efficiency on the macroeconomic level as well. An extensive sensitivity analysis in which especially the errors in the data are analysed is therefore important.

Finally, the simulations provided the insight that for the transition economy of Mongolia – and in all probability also for most other transition economies – the data provide evidence that the closure of enterprises resembles a random way of closure rather than a closure of the least efficient enterprises. This finding sheds a positive light on ideas of some international organisations in Mongolia to reopen closed enterprises in some regions of Mongolia in order to generate employment. This finding contradicts the theory international financial institutions adhere to, that the least efficient enterprises close during a process of transition, and therefore that TFP rises.

The use of TFP for a transition economy in a period of structural economic change, such as Cheng has done in his paper, is not recommended as basically in those situations TFP is an unreliable indicator. With regard to Mongolia, both the analysis of TFP and the use of simulations shed a serious doubt on the conclusion of Cheng that efficiency in Mongolia increased during the first decade of transition. With regard to the method of calculating TFP, this chapter provides evidence that on an ordinal scale in many commonly occurring cases there is no need to estimate $\alpha$ or use regression analysis in order to calculate TFP and to know whether TFP increases or decreases. The authors, furthermore, recommend doing a sensitivity analysis when doing a regression analysis and using the Cobb-Douglas function to calculate TFP, and have therefore extended the method with a sensitivity analysis. This chapter finally stresses the importance of high quality data for the use of an indicator such as TFP. When data are of a lesser quality, there may be a serious flaw in the calculation, and thus the value, of TFP.

## *Appendix A4: Proof of Theorem 1*

THEOREM 1. *A function g on* $\mathrm{R}_{+}^{2}$, *with existing partial derivatives in* $x_1$, $x_2 > 0$, *where, in all applications,* $g(x_1, x_2)$ *denotes GDP, and* $x_1$ *and* $x_2$ *denote capital and labour inputs measured on ratio scales, is uniquely described by*

$$g(x_1, x_2) = \Phi\left(\ln\left(x_1^{\,a} x_2^{\,b}\right)\right), \tag{32}$$

*where a, b are real-valued constants and* Φ *is an exponential function*.

PROOF. Consider a point $(m_1, m_2)$, such that $m_1, m_2 > 0$. The vector $\nabla g(m_1, m_2)$, given by

$$\left(\frac{\partial g(m_1, m_2)}{\partial x_1}, \frac{\partial g(m_1, m_2)}{\partial x_2}\right), \tag{33}$$

is both a gradient of $g$ in $(m_1, m_2)$ and of the line with equation

$$\frac{\partial g(m_1, m_2)}{\partial x_1}(x_1 - m_1) + \frac{\partial g(m_1, m_2)}{\partial x_2}(x_2 - m_2) = 0. \tag{34}$$

Expression (34) is a tangent line of the set $\{(x_1, x_2): g(x_1, x_2) = g(m_1, m_2)\}$ (i.e., the set of indifference points of $g$, which all have the value $g(m_1, m_2)$). This property of $g$ must be invariant under all similarity transformations $\phi_i$, which transform $(m_1, m_2)$ into $(a_1m_1, a_2m_2)$. Following the above arguments, the vector $\nabla g(a_1m_1, a_2m_2)$ is both a gradient of $g$ in $(a_1m_1, a_2m_2)$ and of the line

$$\frac{\partial g(a_1m_1, a_2m_2)}{\partial x_1}(x_1 - a_1m_1) + \frac{\partial g(a_1m_1, a_2m_2)}{\partial x_2}(x_2 - a_2m_2) = 0, \tag{35}$$

which is a tangent line of the indifference set $\{(x_1, x_2): g(x_1, x_2) = g(a_1m_1, a_2m_2)\}$. The ratio scale transformations and the aforementioned invariance condition imply the following relations between the coefficients of equations (34) and (35):

$$\frac{\partial g(a_1m_1, a_2m_2)}{\partial x_1} = \frac{1}{a_1}\frac{\partial g(m_1, m_2)}{\partial x_1}, \tag{36}$$

$$\frac{\partial g(a_1m_1, a_2m_2)}{\partial x_2} = \frac{1}{a_2}\frac{\partial g(m_1, m_2)}{\partial x_2}. \tag{37}$$

Let us introduce the shorthand notation

$$s(m_1, m_2) = \frac{\partial g(m_1, m_2) / \partial x_1}{\partial g(m_1, m_2) / \partial x_2}. \tag{38}$$

By dividing both sides of expressions (36) and (37) we obtain the following functional equation for *s*:

$$s(a_1 m_1, a_2 m_2) = \frac{a_2}{a_1} s(m_1, m_2). \tag{39}$$

This equation can be solved by putting $m_1 = m_2 = 1$, which gives:

$$s(a_1, a_2) = \frac{a_2}{a_1} s(1,1). \tag{40}$$

Since $s(1, 1)$ is a constant and $a_1$, $a_2 > 0$ are arbitrary, expression (40) is the general form of *s*. Let us write $s(1, 1)$ as $a/b$, where *a* and *b* are real-valued constants. Since (38) holds for all $m_1$, $m_2 > 0$, expressions (38) and (40) give rise to the first-order partial differential equation

$$bx_1 \frac{\partial g}{\partial x_1} - ax_2 \frac{\partial g}{\partial x_2} = 0, \tag{41}$$

which has the general solution

$$g(x_1, x_2) = \Phi\left(\ln\left(x_1^{\,a} x_2^{\,b}\right)\right), \tag{42}$$

where $\Phi$ is an arbitrary function (Polyanin et al., 2002). In order to define *g* for zero inputs $x_1$ and $x_2$ (i.e., capital and labour), we have to choose $\Phi$ to be an exponential function.