UNIVERSITÀ POLITECNICA DELLE MARCHE

DIPARTIMENTO DI ECONOMIA

ON THE DECLINE OF AGRICULTURE

Evidence from Italian Regions in the Post-WWII Period

ROBERTO ESPOSTI

QUADERNO DI RICERCA n. 300

Ottobre 2007

ON THE DECLINE OF AGRICULTURE

Evidence from Italian Regions in the Post-WWII Period

Roberto Esposti[§]

Abstract

This article investigates the long-run decline of the agricultural sector during economic development and its contribution to this process. A two-sector model is proposed where the share of agriculture, relative prices and capital accumulation are simultaneously determined within the economy. Under this general equilibrium framework, short-run adjustments with respect to long-run equilibria are admitted through a set of alternative dynamic specifications. The model is then applied to the panel dataset of 20 Italian regions observed over the period 1951-2002 of dramatic economic development but still persistent disparities between Southern and Northern regions.

Keywords: Structural Change, Agriculture, Panel Data, SVAR Models
Running head: On the decline of agriculture
EconLit Classification: O110, Q100, C330

[§] *Department of Economics, Università Politecnica delle Marche (Ancona – Italy), Piazza Martelli, 8 - 60121 Ancona – Italy. Tel. +39 071 2207119, Fax +39 071 2207102, e-mail: r.esposti@univpm.it, website: www.dea.unian.it/esposti.*

The author wishes to thank Franco Sotte, Andrea Bonfiglio, Antonello Lobianco for their helpful suggestions and comments on an earlier version of this paper. The usual disclaimer applies

Contents

On the Decline of Agriculture

Evidence from Italian Regions in the Post-WWII Period

1. Introduction

The complex relation between agriculture and economic growth is a classic topic of Development Economics. After the seminal contributions in '50s and '60s (of which Lewis, 1954, is unanimously considered the initiator), however, the interest on the subject in both theoretical and applied research experienced a long period of decline. But since late '80s, researchers have renewed the interest on this topic with several empirical contributions, even in recent years (Tiffin and Irz, 2006, Sun et al., 2007; Self and Grabowski, 2007).

The basic questions underlying this more recent literature, however, remain the same (Anderson, 1987): Which is the contribution of agriculture to economic growth? Which are the "cause and effect" relations between agricultural and economic growth? Why agriculture declines with economic growth and will this process never arrest or even invert? Which are the driving forces of this decline?

This renewed interest may be explained on several grounds. First of all, since late '80s the so-called new growth theory and the consequent increasing number of empirical contributions on the causes of growth and on growth convergence, have reignited the debate on the role of sectoral composition and structural change within growing economies (Timmer, 2002; Diao et al., 2007). Secondly, the role of agriculture is nowadays again in the foreground with regard to the very rapid, but somehow controversial, economic growth of major developing countries such as China and India, as well as to the growth stagnation of most African countries.

The major novelty in this recent literature, however, is the increasing number of empirical contributions. Although the theoretical hypotheses and models on which they are based can be traced back to '50s and '60s, in fact, such a rich empirical literature only emerged in the '90s, when long time-series combined with large cross-section data (thus, large panels) and appropriate econometric procedures allowed for more sophisticated approaches. Gemmell et al.(2000), Martin and Warr (1993, 1994), Punyasavatsut and Coxhead (2002), Tiffin and Irz (2006), Sun et al. (2007), just to mention a few, are examples of how far these empirical studies can take the analysis with respect to the traditional approaches (Anderson, 1987; Syrquin, 1988).

Most empirical studies actually focus on less developed or developing countries, where agriculture is, in many respects, still the most important sector within the economy; applications to developed "Western" countries are much less frequent. Nonetheless, long time series are now available for several developed countries, and this would allow analysing the whole process of agricultural decline during long-term economic development, going from the stage of underdeveloped and subsistence agrarian economies to post-industrial countries.

This is the case of Italy, whose intense industrial development is relatively recent (mostly occurred in the post-WWII period) but, at the same time, quite heterogeneous across its territories. Although all the 20 Italian regions experienced a dramatic economic growth particularly from '50s to '70s, yet they showed very diverse initial conditions, from already industrialised Northern regions to agrarian Southern ones, and these large disparities did not vanish even after several decades of intense and regular economic growth.

Observing the 20 Italian regions over the 50 years between the WWII and the new century, we can thus observe both cases similar to the currently developing countries (Italian Southern regions in early '50s) and cases analogous to currently most developed countries (Italian Northern regions in recent years). This wide range of diverse conditions can be of major interest in understanding the role of agriculture in economic growth, as well as the major forces driving its decline, according to region-specific initial conditions and characters.

This paper analyses the agricultural decline in Italian regions during the 1951-2002 period. Section 2 reviews the theoretical and empirical literature on this topic, emphasizing its major novelties as well as shortcomings, and the possible improvements in this respect. Section 3 describes in detail the theoretical model here adopted and its main advantages in terms of a more complete representation of the whole set of relations agriculture maintains with non-agricultural sectors.

Section 4 describes how the three equations of the model are empirically specified and the major issues in this respect. In particular, a more realistic description requires a dynamic specification of model equations, allowing for short-run adjustments and disequilibria together with long-run equilibrium relations. Section 5 then presents the major econometric issues implied by the estimation of this dynamic simultaneous equation model when a panel specification is adopted. The estimation strategy, the consequent possible alternative solutions and identification issues are then described.

Section 6 details the data sources and how model variables have been constructed over the whole panel dataset. Section 7 presents model estimates according to the different alternative specifications and estimators, and main results concerning the decomposition and explanation of agricultural decline are critically discussed and compared to previous empirical evidence. Section 8 concludes.

2. Overview of the Literature

We can distinguish two major streams of literature on the decline of agriculture during economic growth. On the one hand, a long and noble tradition of theoretical models trying to represent, in a two-sector environment, how agriculture contributes to economic growth and, in turn, how economic growth affects agricultural development. These models eventually generate plausible and articulated explanations of the driving forces inducing long-term agricultural decline within the economy (as GDP or employment share), though they neither estimate the magnitude of these forces nor demonstrate which of them, in fact, prevails.

On the other hand, a more recent stream of studies tackles this latter aspect more directly, aiming at econometrically estimating the contribution of these driving forces to agricultural decline and, consequently, decomposing it in its basic effects. In doing this and to achieve empirical tractability, however, these empirical studies often lack to fully represent that whole set of complex relations and interdependent equilibria among sectors that eventually generates this decline. Here we shortly summarize the main characters of these two directions of study also aiming at a viable way to combine them in empirical research.

The tradition of two-sector (or multiple-sector) models formalizing the relation between agriculture and economic growth can be traced back to the seminal contribution of Lewis (1954). Lewis'model then generated a first wave of contributions (mainly theoretical) in late '50s and '60s trying to elaborate, and even contradict, Lewis's conclusions towards increasing realism and complexity in the representation of linkages among sectors and of structural change.

Reviewing this literature is well beyond the scope is this paper (see Timmer, 1988 and 2002, for a more extensive treatment, and Gardner, 2003, for a more specific agricultural perspective). Nonetheless, it may be noticed that one major issue emerging in this first wave of studies was the integration of these models within the one-sector neoclassic growth theory upcoming in those same years (Solow, 1956). Attempts were made to adapt the neoclassical growth model to a two-sector context trying to reproduce the structural transformation accompanying economic growth (Uzawa, 1961, 1963, 1964; Takayama, 1965). These attempts soon revealed that Lewis-type models could be hardly reconciled with multisector General Equilibrium (GE) representations of the economy.[1] These theoretical complications, together with the progressive loss of interest in growth and development theory in '70s and early '80s, motivated the end of this first wave of development models.

[1] In particular, the Lewis model strongly relies on inequalities and asymmetries between the two sectors, while GE models are based on the assumption of always-equalized factor prices and marginal productivity (Syrquin, 1988; Lundberg and Squire, 2003).

A second wave, however, appeared in late ‘90s mainly as a consequence of the renewed interest in growth theory as re-designed by the so-called new growth theorists (Timmer, 2002). These contributions deal at once with the stylised facts of (aggregate) growth and of structural change (Syrquin, 1988), that is, change in sector proportions. They are prevalently grounded on a GE framework, namely the neoclassical growth model, in either the Solow or Ramsey version, or more recent new growth models. In this respect, and in chronological order, we can mention Echevarria (1997), Mundlak (2000), Laitner (2000), Kongsamut et al. (2001), Hansen and Prescott (2002), Gollin et al. (2002), Irz and Roe (2005), Foellmi and Zweimuller (2006), Weisdorf (2006), Ngai and Pissarides (2007). Most of these models are two-sector representations of the economy, though some multisector models have also been proposed adopting an I-O or SAM representation (Ahammad, 2002).[2]

These models aiming at combining growth and structural change, however, encounter a major drawback. Though they are strongly stylised GE models and supposedly grounded on sound empirical facts (the so-called “Kaldor facts”), their analytical complexity often makes empirical assessment unaffordable. The replication of these empirical facts is thus obtained as model outcome through calibrations and simulations (see Gollin et al., 2002, as an example). The lack of data, and in particular of long time-series or panel dataset, is also often blamed as one major reason preventing from extending these theoretical approaches to empirical studies (Gemmell et al., 2000).

The inadequate empirical counterpart of this large set of theoretical contributions makes difficult to eventually interpret the role of agriculture in economic growth, as different outcomes emerge, in this respect, from the different models. For instance, Weisdorf (2006) supports the idea of the leading role of non-agricultural (manufacturing) productivity growth with agricultural growth being the follower (this argument seems to be generally accepted even by Gardner, 2003), while in other models agricultural growth precedes and, to some extent, causes aggregate growth and industrial take-off (Irz and Roe, 2005). Interestingly, the empirical assessment of what comes first between agricultural and aggregate economic growth provides mixed results, depending on the country or hystorical context considered (Tiffin and Irz, 2006).

At the end, the only regularity for which no counterfactual cases may be found, and on which any theoretical model is expected to convergence, is the declining share of agriculture

[2] An exhaustive overview of these recent contributions can be found in Greenwood and Seshadri (2001), Robinson (2001) and Timmer (2002).

within the economy.[3] Also in this respect, however, the abovementioned models provide ambiguous outcomes. Three major driving forces, in fact, contribute to this decline. The Engel's law should eventually operate through a declining terms of trade of agricultural products, but also the technological gap (difference in productivity) and different capital intensity (the capital-labour ratio) between the two sectors may be relevant, as well.

In Weisdorf (2006), however, the agricultural decline is apparently explained in combination with an increasing relative price of food products. At the same time, while in some models (and also in actual country experiences) an higher agricultural productivity and capital intensity may favour agricultural decline, in other models and cases decline is accelerated by a relatively lower productivity and capitalization in agriculture. Anderson (1987) provides an extensive explanation of these apparent contradictions as, at least in principle, very different (and somehow opposite) initial conditions under specific circumstances (for instance, open or closed economies) may generate the same outcome in terms of agricultural decline.

Thus, while the decline of agriculture is an indisputable stylised fact of growth, its explanation becomes an empirical question, and this, together with the increasing availability of long time series and large panels, may explain the advent, in the last 15 years, of the second stream of research mentioned earlier in this section. Though using simple quantitative tools, Anderson (1987) may be somehow considered the forerunner of this research stream. Then, a sequence of econometric studies proposed more sophisticated approaches to decompose the agricultural decline in several complementary effects. In strict chronological order, we can mention Martin and Warr (1993, 1994), Gemmell et al. (2000), Punyasavatsut and Coxhead (2002), Sun et al. (2007).

In all these empirical applications, the interaction among sectors is limited to the output supply side, while product demand and input supply are usually disregarded. In this sense, these approaches actually estimate sort of "reduced" forms of the GE two-sector models mentioned above.[4] The three major forces driving agricultural decline enter as exogenous effects while, indeed, they are themselves induced by the development process.

---

[3] Syrquin (1988) provides an exhaustive review of all stylised facts accompanying agricultural decline during economic growth.

[4] It must be underlined, in fact, that the modelling approach followed here, as well as in previous analogous empirical studies (Martin and Warr, 1993 and 1994; Punyasavatsut and Coxhead, 2002), analyses structural change, within a long-run general equilibrium context, that is assuming that factors' marginal productivity, as well as factors' prices, is always equal across sectors, and factors freely move among sectors (this is actually implied by the use of the Production Possibility Frontier). Thus, there is a substantial difference with respect to those models (of which Lewis, 2004, is the most known example) where sectoral decline is explained by a lower factor (namely, labour) productivity, therefore wage differential, pushing labour force out from agriculture towards non-agricultural sectors.

The present paper aims at estimating an econometric model where agricultural share is simultaneously determined together with its major causes, that is, the terms of trade (i.e, relative price) of agriculture with respect to the rest of the economy, the different capital intensity and technological level across the two sectors. These two latter aspects eventually generate different labour productivity.[5] Therefore, the paper also aims at finding an ideal compromise between the two abovementioned streams of research, thus reducing the distance between empirically unworkable two-sector growth models and empirical contributions implicitly based on reduced-form representations of the economy-wide relations.[6] To achieve this, a dynamic panel specification is adopted to allow for a common representation of the underlying theoretical equilibria but still admitting short-run adjustments and disequilibria, as well as structural differences across regions.

This latter aspect is of particular relevance here given the different growth stage of Italian regions over these 50 years. Although the contribution of agriculture to Italian post-war economic development has been largely emphasized, as well as the forces driving the consequent agricultural decline (Fuà, 1992), there are surprisingly few empirical works assessing the actual combination of these forces and the region-specific patterns eventually generating the aggregate result. In fact, as parts of the same country, differences among regions can not be attributed to different fiscal, monetary or sectoral policies. As a matter of fact, these differences can be only explained as different intensity and combination of the abovementioned driving forces.

3. The Model

The main objective of this section is to define a model that can be viewed as the empirical counterpart of two-sectors GE growth models but with specific attention on agriculture. For this major reason, here we model either the supply and the demand side of the economy, also including capital accumulation process and public-government expenditure.

Following Anderson (1987) and Mundlak (2000), we assume a two-sector economy (A=agriculture; E=rest of the economy), producing two goods with quantities Y_A and Y_E, and prices P_A and P_E, respectively. Both products are obtained using two inputs, labour (L) and

---

5 In fact, many authors (see Anderson 1987, for instance) consider relative price change and difference in labour productivity as the major explanations of the decline of agricultural share. Nonetheless, more correctly, labour productivity gap can be either explained in terms of different capital intensity or different technological level.

6 In some empirical papers, for instance Martin and Warr (1994) and Punyasavatsut and Coxhead (2002), the endogeneity of these effects with respect to agricultural share is explicitly acknowledged in adopting appropriate econometric procedures (both works estimate a dynamic simultaneous equation model), but it is never given a clear theoretical justification.

capital (K), although possibly with a different intensities and, thus, factor proportions.[7] The optimal (equilibrium) quantities and prices of both goods are simultaneously determined within the economy as depicted in Figure 1.

Given an initial aggregate factor endowment, L^* and K^*, the Production Possibility Frontier (PPF) defines the whole set of possible efficient combination of Y_A and Y_E. The shape of the PPF depends on the production technology underlying the two sectors. Among these combinations, the actual relative price P_A^*/P_E^* identifies the optimal production mix, Y_A^* and Y_E^*; consequently, the sectoral shares within the economy: $S_A^* = P_A^* Y_A^* / \left(P_A^* Y_A^* + P_E^* Y_E^*\right)$ and $S_E^* = 1 - S_A^*$. The relative price P_A^*/P_E^*, however, is itself endogenous as it is identified by the tangent point between the PPF and the consumption Indifference Curve (IC), whose shape depends, in turn, on the underlying consumers' utility (U^*), that is, preferences.

In Figure 1, economic growth is represented as a forward shift of the PPF generated by factor accumulation (L' and K') and/or technological progress, as well as a forward shift of the IC (U'), due to higher income ($\Delta \mathbf{P'Y}$). Non-homothetic preferences are assumed, thus implying a non-linear Income Expansion Path (IEP) (also known as Income Offer Curve), which is the curve of all optimal combinations of outputs as income grows. It follows that economic growth implies a new optimal mix, Y_A' and Y_E', a new relative price, P_A'/P_E', and, eventually, new sectoral shares, S_A' and S_E'. The shape of the IEP actually expresses the change in sectoral shares, that is, which sector is indeed declining. From this curve, we can also derive the direct relation between income and consumption level or share (the so-called Engel Curve and Scale Curve, respectively) (Matsuda, 2005).[8]

Figure 1 shows how the aggregate supply and demand sides of the economy find an equilibrium that makes both quantities and (relative) prices, thus sectoral shares, endogenous. This representation also makes explicit all the three traditional driving forces of agricultural decline (Martin and Warr, 1993, 1994; Coxhead and Punyasavatsut, 2002; Sun et al., 2007):

---

7 Land (T) and human capital (H) could be also included in this specification as further factors of production as done, for instance, by Coxhead and Punyasavatsut (2002). Here, T is included in the K input, even because it remained almost constant over the whole period in most Italian regions, while H is excluded due to problems in finding appropriate data for this variable over time and across regions and sectors (see Maietta, 2004, for more details about the contribution of H to the Italian agricultural development in the last decades).

8 Analogously, the change in consumption level due to income increase is called *Engel's effect*, while the change in consumption share is the *Scale effect*.

the Engel's law (through change in relative price, P_A / P_E), the Rybczynski Theorem[9] (through change in capital intensity, K/L), and the technological gap (through change in the shape of PPF due to different technical change rates across sectors). These forces involve both demand and supply, and reciprocally interact and influence each other.

To make this representation fully consistent with a general equilibrium model, we also need to endogenize the change in input supply and, possibly, technical change itself. To achieve this, we assume that Figure 1 represents open regional economies (Mundlak, 2000; pag. 37) where openness is, indeed, limited to trade (non-zero net exports are admitted for both goods), while the economy is closed in terms of net savings, since all regional investments are funded by domestic net savings. Migration is admitted but only in the sense that the supply of labour (i.e., L endowment) is exogenous, so it incorporates both natural and migration regional balance.

The model here adopted thus formalizes the simultaneous relations underlying Figure 1. Duality reveals particularly useful for this kind of representation by means of the *revenue function*, generating the *supply-side share equation* on the supply side, and of the *distance function*, generating the *inverse uncompensated demand function* on the demand side. Next sections provide the description of the whole model structure and the respective empirical functional specifications (section 4).

3.1. Supply Side

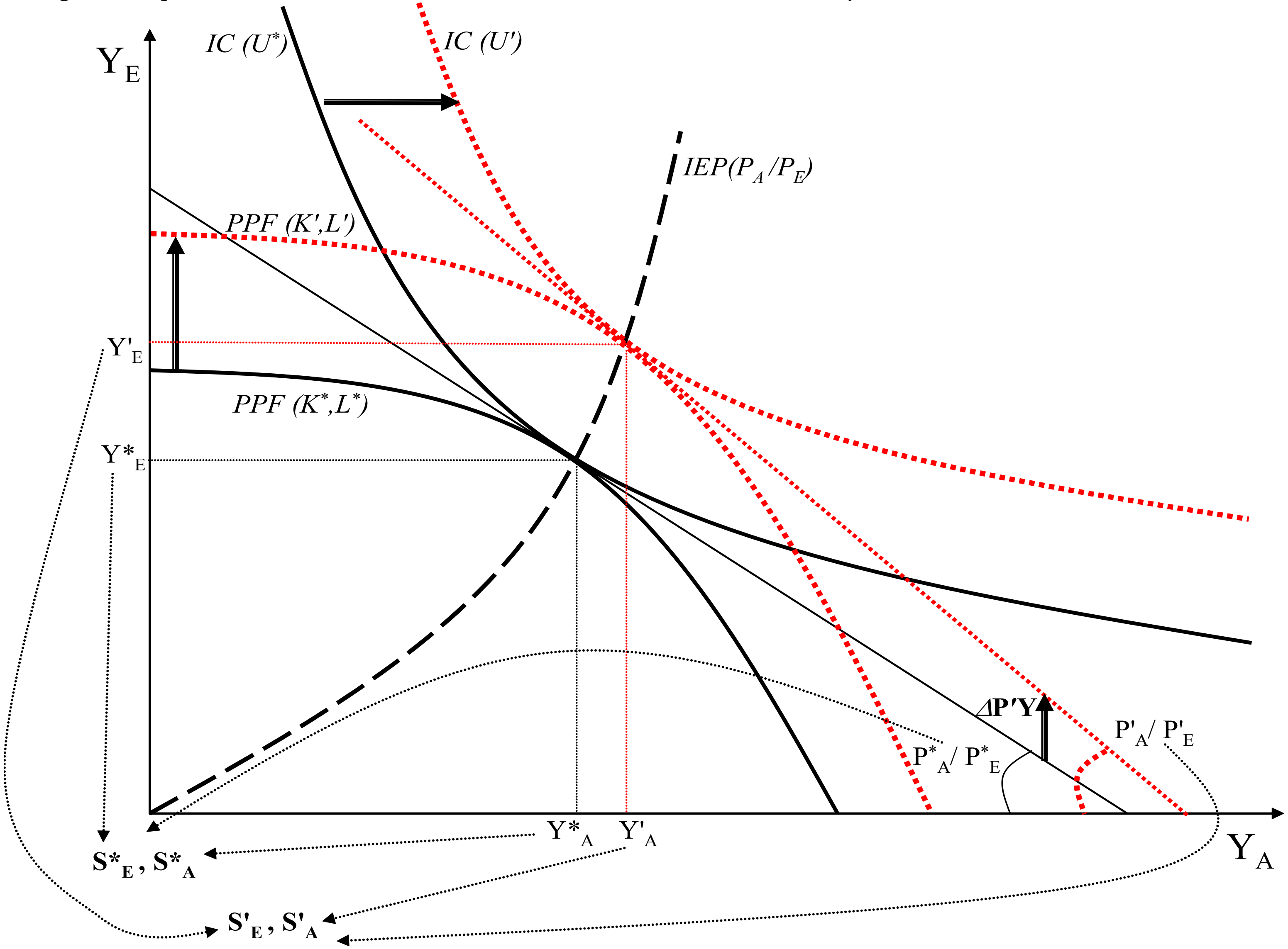
3.1.1. Technology: PPF and Share Equations

Consider a small open regional economy (in the sense outlined above) and assumes it produces with constant returns to scale (overall and in all sectors) and competitive equilibria in all markets (inputs and outputs). Given factor endowments and output prices, the optimal production mix between the two goods is found on the technological frontier described by the implicit Production Possibility Frontier (PPF):

(1) $F(Y_A, Y_E, K, L) = 0$

9 This theorem states that "the accumulation of capital *per* worker, as economic growth proceeds, tends to encourage expansion of the more capital-intensive sector at the expense of the more labour-intensive sector" (Anderson, 1998, p. 198). In the present case, by Rybczynski effect we mean the effect of increasing aggregate capital intensity on agricultural share: a negative effect would imply that agriculture is more labour-intensive.

Figure 1. Equilibrium between A and E: Y_A, Y_E and P_A/P_E are simultaneously determined



Source: our adaptation from Anderson (1987, p. 197) and Mundlak (2000, p. 35)

The optimal allocation of K and L across sectors, thus the optimal mix (Y_A^* and Y_E^*), must maximize the revenue:[10]

$$\text{(2a)}\quad r(\mathbf{P},\mathbf{Z})=\max_{K_A,L_A}\left[\mathbf{P}\cdot\mathbf{Y}:\ \mathbf{Y}\in\mathbf{Y}(\mathbf{Z})\right],\qquad \mathbf{P},\mathbf{Y}\in\Re^2;\ \mathbf{Z}\in\Re^2$$

where $\mathbf{Y}(\mathbf{Z})$ is the convex production set given factor endowment $\mathbf{Z}=\{K,L\}$; $\mathbf{P}=\{P_A,P_E\}$ and $\mathbf{Y}=\{Y_A,Y_E\}$. Thus, on the output supply side, prices and factor endowments are exogenous, while the sectoral outputs and shares are endogenously determined.

$r(\mathbf{P},\mathbf{Z})$ is the *revenue function* (Chambers, 1988; Coxhead and Punyasavatsut, 2002; Martin and Warr, 1993, 1994); when it explicitly refers to aggregate economies, it is also called *GDP function* (Kholi, 1978, 1991; Sun, et al., 2007). This dual function has the following theoretical properties: it is positive, continuous, non-decreasing in $\mathbf{P}$ and $\mathbf{Z}$; linearly homogeneous and concave with respect to $\mathbf{P}$ (Chambers, 1988; Kholi, 1978, 1991).

Under these regularity conditions, we can apply the Hotelling lemma, thus obtaining the following supply functions, $Y_i(\mathbf{P},\mathbf{Z})$:

$$\text{(3)}\quad Y_i(\mathbf{P},\mathbf{Z})=Y_i^*=\frac{\partial r(\mathbf{P},\mathbf{Z})}{\partial P_i},\qquad i=A,E$$

This relation remains valid up to a monotonic transformation: by expressing the revenue function in logarithmic form, the application of the Hotelling Lemma gives the *sectoral share equation*:

$$\text{(4a)}\quad S_i(\mathbf{P},\mathbf{Z})=\frac{\partial \log r(\mathbf{P},\mathbf{Z})}{\partial \log P_i}=\frac{\partial r(\mathbf{P},\mathbf{Z})}{\partial P_i}\frac{P_i}{r}=\frac{Y_i\,P_i}{r},\qquad i=A,E$$

where $\sum_i S_i(\mathbf{P},\mathbf{Z})=1$, $i=A,E$. The properties of this share equation are directly and easily derived from the abovementioned regularity conditions of the revenue function.

3.1.2. Sectoral Technological Change

Given the amount of production factors, L^* and K^*, the PPF may shift forward only if technical change occurs in one or both sectors. Here, we allow technological level to differ across sectors. T_A and T_E indicate technological levels in agriculture and the rest of the economy, respectively, and we assume both exogenous. T_E is proxied with the usual time trend t, i.e. $T_E=e^{\lambda_E t}$; thus, technological change rate in E is constant over time:

[10] The argmax of problem (2) could be more intuitively written as Y_A^* and Y_E^*. However, the real underlying problem, given the sectoral production functions, concerns the allocation of production factors among the two sectors, i.e. L_A and K_A that univocally determine the amount of inputs used in the other sector, L_E and K_E.

$\overset{\bullet}{T}_E = \frac{\partial \ln T_E}{\partial t} = \lambda_E$. T_A is actually determined by two components (Esposti and Pierani, 2000, 2006). One is generated by agricultural public R&D (which is related to the overall public expenditure - see section 3.2.3), R_A ; the other component is analogous to T_E being proxied by the conventional time trend. Thus, $T_A = e^{\lambda_A t} R_A^{\alpha_1}$, with $\lambda_A \neq \lambda_E$. Consequently, technical change rate in agriculture may differ from the other sector and can vary over time as effect of changes in public R&D growth rate: $\overset{\bullet}{T}_A = \lambda_A + \alpha_1 \overset{\bullet}{R}_A$.

Upon the introduction of sectoral technological levels, the revenue function can be generalized as follows:

(2b) $r(\mathbf{P}, \mathbf{Z}, R_A, t) = \max\limits_{K_A, L_A} \left[\mathbf{P} \cdot \mathbf{Y} : \ \mathbf{Y} \in \mathbf{Y}(\mathbf{Z}, R_A, t) \right], \qquad \mathbf{P}, \mathbf{Y} \in \Re^2; \mathbf{Z} \in \Re^2$

and the sectoral share equation becomes:

(4b) $S_i(\mathbf{P}, \mathbf{Z}, R_A, t) = \frac{\partial \log r(\mathbf{P}, \mathbf{Z}, R_A, t)}{\partial \log P_i} = \frac{\partial r(\mathbf{P}, \mathbf{Z}, R_A, t)}{\partial P_i} \frac{P_i}{r} = \frac{Y_i P_i}{r} \quad , \qquad i = A, E$

According to this modelling approach, variables K, L, P_A, P_E, R_A and, obviously, t, are exogenous on the output supply side. Most of these variables, however, become endogenous once the demand and input supply sides are also modelled; in practice, only L, R_A and t will remain exogenous.[11]

3.2. Demand Side and Capital Supply

3.2.1. Preferences: Distance Function and Inverse Demand Function

In modelling consumers' preferences, the main purpose here is to make relative prices endogenous, that is, dependent on sectoral shares. Therefore, following Mundlak (2000), Deaton and Mullenbauer (1980) and, above all, Wong and McLaren (2005), we model the demand side through *Inverse Demand Functions*.

Preferences can be described by a direct utility function, $u = U(\mathbf{X})$, where $\mathbf{X} = \{X_A, X_E\}$ is the consumption vector. The aggregate consumers' behaviour can be thus obtained by maximizing this utility function under their respective budget constraint. Upon the

---

[11] Modelling labour supply could be also possible to better take into account the important role of net migration across regions (mainly from South to North) in the early decades of the period under study. This process also strongly affected agriculture, favouring the loss of labour force. Here, however, we do not pay specific attention to this aspect, because it would imply a careful evaluation of labour quality, that is human capital and age structure, which is, as mentioned, a problematic issues beyond the scope of this study. However, improvements in this direction could be interesting in future research.

specification of the direct utility function $U(\mathbf{X})$, we can apply the Hotelling-Wold identity to obtain the inverse uncompensated demand (Kim, 1997):

$$(5)\quad R_i(\mathbf{X}) = \frac{\partial U(\mathbf{X})}{\partial X_i} \Bigg/ \left(\sum_i \frac{\partial U(\mathbf{X})}{\partial X_i} X_i \right), \qquad i = A, E$$

where R_i indicates the normalized prices, that is $R_i = \frac{P_i}{C} = \frac{P_i}{\sum_i P_i X_i}$, $i = A, E$, and $C = \sum_i P_i X_i$ is the total value of consumption. Thus, we can also write:

$$(6)\quad \frac{R_i(\mathbf{X})}{R_j(\mathbf{X})} = \frac{P_i}{P_j} = \frac{\frac{\partial U(\mathbf{X})}{\partial X_i} \Big/ \left(\sum_i \frac{\partial U(\mathbf{X})}{\partial X_i} X_i \right)}{\frac{\partial U(\mathbf{X})}{\partial X_j} \Big/ \left(\sum_i \frac{\partial U(\mathbf{X})}{\partial X_i} X_i \right)}$$

Instead of using the utility function directly, we can also describe preferences by means of a dual function, the *distance function* $d(\mathbf{X},u)$, defined as:

$$(7)\quad d(\mathbf{X},u) = \max_d \{ d > 0 : U(\mathbf{X}/d) \geq u \}$$

This function thus indicates the maximum amount, *d*, by which the consumption bundle $\mathbf{X}$ must be inflated, or deflated, in order to reach the base utility contour defined by u; then, when $u = U(\mathbf{X})$, $d(\mathbf{X},u) = 1$. The distance function has the following theoretical properties: it is positive, continuous, linearly homogeneous, non-decreasing, and concave with respect to $\mathbf{X}$; non-decreasing in u (Deaton and Mullenbauer, 1980; Wong and McLaren, 2005).

The main advantage of using the distance function comes from the application of the Shephard-Hanoch Lemma to obtain the *Inverse Compensated Demand Function*:

$$(8)\quad R_i^C(\mathbf{X},u) = \frac{\partial d(\mathbf{X},u)}{\partial X_i}, \qquad i = A, E$$

Being u unobservable, equation (6) has not a direct empirical applicability. Nonetheless, upon the specification of the distance function, we can impose and invert the relation $d(\mathbf{X},u) = 1$ to replace the unobservable u with $U(\mathbf{X})$, thus obtaining the *Inverse Uncompensated Demand Functions*:

$$(9)\quad R_i(\mathbf{X}) = R_i^C(\mathbf{X}, U(\mathbf{X})) = \frac{\partial d(\mathbf{X}, U(\mathbf{X}))}{\partial X_i}, \qquad i = A, E$$

whose theoretical properties are directly derived from those of the distance function. It also follows that, for any given pair of commodities *i* and *j*, we can obtain the *relative price equation* as follows:

$$(10)\ \frac{R_i(\mathbf{X})}{R_j(\mathbf{X})}=\frac{P_i}{P_j}=\frac{\dfrac{\partial d(\mathbf{X},U(\mathbf{X}))}{\partial X_i}}{\dfrac{\partial d(\mathbf{X},U(\mathbf{X}))}{\partial X_j}}, \qquad i,j=A,E$$

Equations (9) and (10) exactly correspond to equations (5) and (6), respectively, albeit they are obtained from dual functions (direct utility and distance functions). Therefore, it remains always true that $R_i(\mathbf{X})=R_i^C(\mathbf{X},U(\mathbf{X}))$.

3.2.2. Savings, Investments and Capital Accumulation

We model supply of new capital (investments) by imposing a saving identity. In fact, we simply assume that the net flow of savings and investments is null across regions. Thus, within any regional economy, domestic savings correspond to regional gross investments and can be computed as income less consumption and taxes:

$$(11a) \qquad s=\left(\sum_i P_i Y_i-\sum_i P_i X_i\right)\Gamma, \qquad i=A,E$$

where $\mathbf{P}'\mathbf{Y}=\sum_i P_i Y_i$ is total income and $\left(\sum_i P_i Y_i-\sum_i P_i X_i\right)$ is the amount of savings. Γ is the degree of taxation, that is the percentage of savings needed to pay taxes. Here, taxes are entirely used to fund government expenditures and not for public capital investments, as these are already included in the saving rate *s*.

We can elaborate equation (9a) reminding that, for both goods, regional economies are open to external trade and that part of production is actually re-utilized within the regional economy as intermediate goods. Therefore, we can derive the following relation between domestic consumption, *X*, and production, *Y*: $X_i=(\zeta_i-g_i)Y_i$, $\forall i=A,E$, where ζ_i indicates the net export rate and is > (<) 1 when net exports are negative (positive), while g_i is the percentage of i-th product used as intermediate good, therefore it is always ≥0.

Equation (11a) can be rewritten as:

$$(11b) \qquad s=\left[\sum_i P_i Y_i-\sum_i P_i(\zeta_i-g_i)Y_i\right]\Gamma=\left[\sum_i P_i(1+g_i-\zeta_i)Y_i\right]\Gamma, \qquad i=A,E$$

We can use (11b) to model capital accumulation over time. As *s* corresponds to gross investments, and indexing with respect to time, we can express net investments, *i*, as follows:

$$(12a)\ K_t-K_{t-1}=i_t=(1-d)s_t=(1-d)\left[\sum_i P_{it}(1+g_{it}-\zeta_{it})Y_{it}\right]\Gamma_t, \qquad i=A,E$$

where d indicates the capital depreciation rate. Dividing by K_{t-1} and multiplying by L_{t-1}/L_t , after some simple algebraic manipulations, (12a) can be written in terms of capital intesity:

(12b) $$\frac{K_t/L_t}{\left(K_{t-1}/L_{t-1}\right)} = \frac{k_t}{k_{t-1}} = \frac{(1-d)\left[\sum_i P_{it}\left(1+g_{it}-\zeta_{it}\right)Y_{it}\right]\Gamma_t}{L_t\, k_{t-1}} + \frac{L_{t-1}}{L_t}, \quad i = A, E$$

Now consider that, by definition, $S_i = \dfrac{P_i Y_i}{\sum_j P_j Y_j}$; thus, sectoral share ratios can be written as $\dfrac{S_i}{S_j} = \dfrac{P_i Y_i}{P_j Y_j}$. In the two-sector (two-good) case, $S_{jt} = 1 - S_{it}$; this implies $P_i Y_i = P_j Y_j \dfrac{S_i}{\left(1-S_i\right)}$, $i, j = A, E$. It follows that (12b) can be expressed in terms of sectoral shares as follows:[12]

(12c) $$\frac{k_t}{k_{t-1}} = \frac{(1-d)P_{it}\, Y_{it}\left[\left(1+g_{it}-\zeta_{it}\right)+\left(1+g_{jt}-\zeta_{jt}\right)\dfrac{\left(1-S_i\right)}{S_i}\right]\Gamma_t}{L_t\, k_{t-1}} + \frac{L_{t-1}}{L_t}, \quad i = A, E$$

Considering that $L_t = L_{t-1}\left(1+n_t\right)$, where n_t indicates the population growth rate assumed exogenous and equal to employment growth rate,[13] we finally obtain:

(12d)

$$\frac{k_t}{k_{t-1}} = \frac{(1-d)P_{it}\, Y_{it}\left[\left(1+g_{it}-\zeta_{it}\right)+\left(1+g_{jt}-\zeta_{jt}\right)\dfrac{\left(1-S_i\right)}{S_i}\right]\Gamma_t}{L_t\, k_{t-1}} + \frac{1}{\left(1+n_t\right)}, \quad i = A, E$$

3.2.3. Taxes and Public Expenditure

To complete the model, we assume that, in any region, we may observe *deficit spending* in the short-run, that is public expenditure rate G may temporarily exceed the degree of taxation Γ. Therefore: $\Gamma_t = G_t\, u_t$, where $u_t < 1$ $(u_t > 1)$ is the public deficit (surplus) rate. It is

---

12 This result obviously holds only in the two-sector case.

13 This assumption is clearly acceptable in the long-run, when population growth rate inevitably tends to coincide with the employment growth rate. In the short-run, however, they may significantly differ, especially due to change in unemployment rate. Thus, this assumption implies, in the short-run, constant unemployment and participation rates. We could also assume population growth rate constant over time, that is $n_t = n, \forall t$. This is, indeed, a strong assumption as demographic patterns often significantly change as economic development proceeds. Nonetheless, this simplification may be particularly useful in the empirical implementation of the model and should not significantly affect the interpretation of results, as this growth is still exogenous within model itself.

reasonable, on the one hand, to consider this public deficit stochastic, as it depends on unpredictable regional business cycles and on uncontrollable components of public expenditure. On the other hand, this public deficit should be annulled in the long-run. Therefore, years of public deficit should alternate with years of surplus, such that we can treat u_t as a random variable with $E(u_t)=1$ and $Var(u_t)=\sigma_u^2$. It follows that, according to identity (12d), $\frac{k_t}{k_{t-1}}$ itself behaves as a random variable.

We finally assume that public expenditure is made of two components. The first is specific to agriculture, and it is public agricultural R&D; the second is simply a trend. Therefore, we can write $G_t = R_{At}^{\alpha_2}\, e^{\phi t}$ and, taking logarithms, we obtain:

(13) $$\ln \Gamma_t = \alpha_2 \ln R_{At} + \phi t + \ln u_t$$

with $E(\ln u_t)=0$ and $Var(\ln u_t)=\ln\sigma_u^2 = \overline{\sigma}_u^2$. It is also clear, from (13), that the growth rate of the degree of taxation is $\dot{\Gamma}_t = \alpha_2 \dot{R}_{At} + \phi$, and can be either < or > 0, though it is expected to be very close to 0 in the long-run. This also clarifies the economic meaning of parameter α_2. It expresses how much increase in R_{At} is reflected in *G* (and *Γ*) growth, and evidently depends on how large R_{At} is with respect to *i* and, in turn, to $\mathbf{P'Y}$. Therefore, α_2 can be given a sort of "public agricultural R&D intensity" interpretation.

Combining (13) with (12d), we obtain:

(12e)

$$\frac{k_t}{k_{t-1}} = \frac{(1-d)P_{it}\,Y_{it}\left[(1+g_{it}-\zeta_{it})+(1+g_{jt}-\zeta_{jt})\frac{(1-S_i)}{S_i}\right]R_{At}^{\alpha_2}\,e^{\phi t}u_t}{L_t\,k_{t-1}} + \frac{1}{(1+n_t)}, \qquad i=A,E$$

4. Empirical specification

When applied to the two-sector case, the model developed in previous section is made of three equations, (4b), (10) (or (6)), and (12e). To make it empirically tractable, however, some further simplifications are needed. Firstly, we normalize prices by P_E, to express agricultural price in relative terms, $\overline{P}_A = \frac{P_A}{P_E}$. Secondly, to achieve a linear specification of all model equations, logarithmic transformations are used.

4.1. Equations' Specification

4.1.1. Share Equation

Following Punyasavatsut and Coxhead (2002) and Sun et al. (2007), the revenue function (2b) is here specified as a Translog (TL), which is a second-order approximation of a general well-behaved revenue function.[14] The main advantage of the TL, besides flexibility, rests on the derivation of the respective share equations just by taking first-order derivatives. For simplicity, we assume Hicks neutral technical change in both sectors. This assumption prevents the complications occurring within the PPF under biased technological change, also, maintains a clear separation between different technological change rates and Rybczynski effect in analysing sectoral share variations.[15] According to this assumption, in fact, technical change can be simply modelled as a supply function shifter, that is $Y_i(\mathbf{P},\mathbf{Z},T_i)=T_i\,Y_i(\mathbf{P},\mathbf{Z}),\ i=A,E$.

As a consequence, within the revenue function, an Hicks-neutral technical change occurring in the i-th sector (ΔT_i) corresponds to a price increase in the sector itself, as ΔT_i implies higher production value for a given factor use. Therefore, the revenue function can be rewritten as $r(\mathbf{TP},\mathbf{Z})$, where $\mathbf{T}=\{T_A,T_E\}$, that is, $\mathbf{TP}$ is the actual price vector positioning the optimal production mix on the PPF. The TL revenue function is the following (Coxhead and Punyasavatsut, 2002)[16]:

(14)

$$\ln r(\mathbf{TP},\mathbf{Z})=\alpha_0+\sum_i \alpha_i \ln(T_iP_i)+\frac{1}{2}\sum_i \sum_j \alpha_{ij} \ln(T_iP_i)\ln(T_jP_j)+\sum_k \beta_k \ln Z_k$$
$$+\frac{1}{2}\sum_k \sum_h \beta_{kh} \ln Z_k \ln Z_h+\sum_i \sum_k \delta_{ik} \ln(T_iP_i)\ln Z_k,\quad \forall i,j=A,E\ \ and\ \ \forall k,h=K,L$$

from which we can easily derive the agricultural share equation:

(15a) $$S_A=\frac{\partial \ln r(\mathbf{TP},\mathbf{Z})}{\partial \ln(T_AP_A)}=\alpha_A+\alpha_{AE}\ln(T_EP_E)-\alpha_{AE}\ln(T_AP_A)+\delta_{AK}\ln K+\delta_{AL}\ln L$$

---

14 Martin and Warr (1993, 1994) adopt a TL profit function.

15 Assuming Solow-neutral (that is, labour-saving) technical change in both sectors would be also possible, but this would make the definition of the revenue function, and of the consequent supply and share functions, much more complex. On the contrary, assuming Hicks-neutral technical change when actually is not, implies an overestimation of the Rybczynski effect on sectoral shares.

16 Sun et al. (2007) do not impose Hicks-neutral technical change but do not explicitly admit a technological gap among the sectors.

This share equation is homogeneous of degree zero with respect to prices ($\sum_j \alpha_{ij} = 0$), while constant returns to scale are assumed in both sectors.[17] Therefore, we can normalize $T_E P_E$ by $T_A P_A$ and we can express factor quantities in intensive form, that is, by labour unit ($k = \frac{K}{L}$):[18]

(15b) $$S_A = \alpha_A + \alpha_{AE} \ln \frac{P_E}{P_A} + \alpha_{AE} \ln \frac{T_E}{T_A} + \delta_{AK} \ln k$$

According to section 3.1.2. and to the notation introduced before, we can finally write:

(15c) $$S_A = \alpha_A - \alpha_{AE} \ln \overline{P}_A + \alpha_{AE}\, t \left(\lambda_E - \lambda_A\right) - \alpha_{AE} \alpha_1 \ln R_A + \delta_{AK} \ln k$$

This share equation attributes agricultural decline to the combination of four forces: agricultural relative price (taking into account consumption shares, thus the Engel's law); technological gap between agriculture and the rest of the economy, $\left(\lambda_E - \lambda_A\right)$; a declining amount of public resources spent for agricultural R&D to eventually fill this gap; finally, overall capital intensity whose effect on agricultural share depends on the abovementioned Rybczynski effect.

For the Engel's law to hold, α_{AE} is expected to be < 0. Then, also technological gap effect holds true, provided that this gaps exists, that is $\left(\lambda_E - \lambda_A\right) > 0$ and, if $\alpha_1 > 0$, public agricultural R&D really reduces this effect. $\left(\lambda_E - \lambda_A\right)$, however, is not observed, therefore we can only directly estimate $\alpha_{AE} \left(\lambda_E - \lambda_A\right)$, which should evidently be <0. Finally, the Rybczynski effect can be either positive or negative: under Rybczynski effect, δ_{AK} is expected to be <0 if capital intensity is lower in agriculture, and >0 if agricultural capital intensity is higher. This decomposition is similar to previous studies (Martin and Warr 1993, 1994; Coxhead and Punyasavatsut, 2002; Sun et al., 2007). The main difference, here, is that these effects are themselves dependent on agricultural share, that is, are endogenous.

---

17 Constant returns to scale, that is linear homogeneity of the revenue function with respect to input quantities, also implies $\delta_{AK} + \delta_{AL} = 0$, that is, $\delta_{AL} = -\delta_{AK}$.

18 Note that parameters associated to $\ln \frac{P_E}{P_A}$ and $\ln \frac{T_E}{T_A}$ are equal due to the hypothesis of Hick's neutral technical change. Thus, this latter hypothesis can be either imposed, as done here, or empirically tested by simply testing for the equality of the two parameters.

4.1.2. Inverse Demand

To make relative price depend on sectoral share, given that $P_{it} X_{it} = P_{it} Y_{it} (\zeta_{it} - g_{it})$, we can write (9) as $R_i(\mathbf{X}) = R_i(\mathbf{Y}, \boldsymbol{\zeta}, \mathbf{g})$, and (10) as

$$\frac{R_i(\mathbf{Y},\boldsymbol{\zeta},\mathbf{g})}{R_j(\mathbf{Y},\boldsymbol{\zeta},\mathbf{g})} = \frac{P_i}{P_j} = \frac{\dfrac{\partial d(\mathbf{Y},\boldsymbol{\zeta},\mathbf{g},\, U(\mathbf{Y},\boldsymbol{\zeta},\mathbf{g}))}{\partial Y_i}}{\dfrac{\partial d(\mathbf{Y},\boldsymbol{\zeta},\mathbf{g},\, U(\mathbf{Y},\boldsymbol{\zeta},\mathbf{g}))}{\partial Y_j}} = z\left(\frac{Y_i}{Y_j}, \boldsymbol{\zeta}, \mathbf{g}\right) = z(S_i, \boldsymbol{\zeta}, \mathbf{g}).$$

The generic function $z(.)$ to be explicit, we need an utility function simple enough to generate an empirically workable inverse demand function and that admits non-homothetic preferences, as implied by the Engel's law. On the contrary, we are not particularly interested in flexible functional forms in our two-good model,[19] as it seems reasonable to impose substitutability relation: we assume the two goods are substitutes to some extent and, though we do not wish to impose the elasticity of substitution, still we assume it remains constant over time.

Instead of specifying a distance function, however, we derive an empirical specification of equation (5) from some simple direct utility function. It is of major interest, here, the "CES function with subsistence requirements" introduced by Beckman and Smith (1993; see also Primont and Primont, 1995). This function is relatively straightforward in the two-good case and is particularly suitable for the analysis of agricultural or food demand, as it is non-homothetic and can be considered an evolution of the well-known Stone-Geary functional form (Beckman and Smith, 1993).[20]

In the two-good (A, E) case, this CES direct utility function is:

(16) $U(\mathbf{X}) = \left[\delta(X_A - \gamma_A)^{-\rho} + (1-\delta)(X_E - \gamma_E)^{-\rho}\right]^{-1/\rho}$

where $0 \le \delta \le 1$ and $-1 < \rho \ne 0$ are the distribution and substitution parameters, respectively; γ_A and γ_E are the respective minimum requirement levels.

Applying the Hotelling-Wold identity, Beckman and Smith (1993) show that the inverse demand functions implied by the direct utility function (16) are:

(17a) $$R_A = \frac{P_A}{r} = \left[I\,\delta X_E^{(1+\rho)}\right] \Big/ \left[\delta X_A X_E^{(1+\rho)} + (1-\delta) X_E (X_A - \gamma_A)^{(1+\rho)}\right]$$

---

19 There are several alternative flexible specifications of the distance function that allows for the derivation (through Shepard's Lemma) of flexible inverse demand equations (Brown *et al.*, 1995; Wong and McLaren, 2005).

20 We do not consider Cobb-Douglas utility function because it imposes homothetic preferences and elasticity of substitution = 1; for a detailed analysis of the Engel curves (or functions) implied by Stone-Geary utility functions, see Aasness et al. (1993).

(17b) $$R_E = \frac{P_E}{r} = \left[I\left(1-\delta\right)\left(X_A - \gamma_A\right)^{1+\rho} \right] \Big/ \left[\left(1-\delta\right) X_E \left(X_A - \gamma_A\right)^{(1+\rho)} + \delta X_A X_E^{(1+\rho)} \right]$$

where I indicates the disposable nominal income, that is $I = \mathbf{P'Y}$; moreover, we have assumed, for simplicity, that there is no minimum requirement level for good E, that is, $\gamma_E = 0$.[21]

After some simple algebraic manipulations, we obtain an empirical version of (6) in logarithmic form:

(17c) $$\ln \frac{P_A}{P_E} = \ln \overline{P}_A = \ln\left(\frac{\delta}{1-\delta}\right) + \left(1+\rho\right) \ln \frac{X_E}{\left(X_A - \gamma_A\right)}$$

The economic interpretation of parameters in (17c) deserves some further comment. According to Beckman and Smith (1993, p. 299), we can write the elasticity of substitution as $\sigma = d \ln\left[\left(X_A - \gamma_A\right) / \left(X_E - \gamma_E\right)\right] / d \ln\left(P_A / P_E\right) = 1/\left(1+\rho\right)$; thus, it exclusively depends on ρ, while the scale elasticities (or scale effects) are derived for both goods in Appendix 1. The distribution parameter, δ, is involved in the calculation of the demand income elasticity. This elasticity identifies the scale effects, that is, how consumption proportion between the two goods changes as the overall consumption level increases. As shown in Appendix 1, since utility function (16) implies non-homothetic preferences, income elasticity differs between the two goods and in both cases is $\neq 1$. In particular, regardless γ_A, δ and ρ, good A is always a non-superior good while good E is a luxury good, being income elasticity always < and > 1, respectively.

(17c) and (15c) are simultaneous equations. Simultaneity emerges clearly whenever we express (17c) in terms of S_A. To achieve this, we have, firstly, to express $X_{A,E}$ in terms of $Y_{A,E}$. As already discussed, we can write $X_{A,E} = \left(\zeta_{A,E} - g_{A,E}\right) Y_{A,E}$. If we assume, as seems reasonable, that net exports do not involve the minimum requirement level γ_A and this minimum level is only used for direct human consumption and not as intermediate good, it follows that equation (17c) can we rewritten as:

(17d) $$\ln \overline{P}_A = \ln\left(\frac{\delta}{1-\delta}\right) + \left(1+\rho\right) \ln \frac{\left(\zeta_E - g_E\right)\left(Y_E\right)}{\left(\zeta_A - g_A\right)\left(Y_A - \gamma_A\right)}$$

---

[21] This assumption seems acceptable and realistic considering that, as explained in Appendix 1, good E behaves as a luxury good; therefore, it is reasonable to assume that there is no minimum consumption level to be fulfilled.

Given that, by definition, $S_A = \frac{P_A Y_A}{P_E Y_E + P_A Y_A}$ and $\frac{S_A}{S_E} = \frac{P_A Y_A}{P_E Y_E}$, we can write $\frac{Y_A}{Y_E} = \overline{P}_A \frac{S_A}{S_E} = \overline{P}_A \frac{S_A}{(1-S_A)}$. Multiplying both sides by $\frac{1}{(1-\gamma_A/Y_A)}$, we can write $\left[\frac{(Y_A-\gamma_A)}{Y_E}\right] = \left[\overline{P}_A \frac{S_A}{(1-S_A)} \frac{Y_A}{(Y_A-\gamma_A)}\right]$. Taking logarithms, it becomes $\ln\left[\frac{(Y_A-\gamma_A)}{Y_E}\right] = \ln \overline{P}_A + \ln\frac{S_A}{(1-S_A)} + \ln\frac{Y_A}{(Y_A-\gamma_A)}$. By substitution in (17d) and after few algebraic passages, we finally obtain:

(17e) $\ln \overline{P}_A = a + b\ln\frac{S_A}{(1-S_A)} + b\ln\frac{Y_A}{(Y_A-\gamma_A)}$

where $a = \frac{\ln\left(\frac{\delta}{1-\delta}\right) + (1+\rho)\ln\frac{(\zeta_E - g_E)}{(\zeta_A - g_A)}}{(2+\rho)}$ and $b = -\frac{(1+\rho)}{(2+\rho)}$.

To make (17e) empirically tractable, we need two further approximations. Firstly, we have to approximate $\frac{Y_A}{(Y_A-\gamma_A)}$, as γ_A is unknown and unobserved; then, we need a linear function of S_A, $f(S_A)$, such that $\ln\frac{S_A}{(1-S_A)} \cong f(S_A)$.

By definition of minimum requirement and having assumed that the minimum consumption level is not traded, we may notice that $\frac{Y_A}{(Y_A-\gamma_A)} > 1$ and it tends to 1 as Y_A goes to ∞. Therefore, if we assume that Y_A grows over time at a constant rate, which seems plausible in a long-run perspective, we can approximate this term with a monotonically decreasing function going to 1 as t goes to ∞, that is: $\frac{Y_A}{(Y_A-\gamma_A)} \cong (\beta)^{1/t}$, where $\beta \geq 1$ is an unobserved term (γ_A being unobserved) defined as the initial value (t = 0) of the ratio to be approximated: $\frac{Y_{A0}}{(Y_{A0}-\gamma_A)} = \beta$.

With respect to term $\ln\frac{S_A}{(1-S_A)}$, we may notice that $0 \leq S_A \leq 1$ and, at least in the regional sample adopted here, it is also always $0{,}01 \leq S_A \leq 0{,}40$. Within this interval, it can be shown that $\ln\frac{S_A}{(1-S_A)}$ is well approximated by a cubic function, that is:

$\ln\frac{S_A}{(1-S_A)}=c_0+c_1S_A+c_2S_A^2+c_3S_A^3+\tilde{u}$, where $\tilde{u}$ is a time-varying error term generated as rest of the approximation. Appendix 2 provides more details on this approximation and on value of parameters c_0, c_1, c_2, c_3.

Consequently, equation (17e) can be rewritten as follows:

(17f) $$\ln\overline{P}_A=(a+bc_0)+bc_1S_A+bc_2S_A^2+bc_3S_A^3-bt\ln\beta+\overline{u}$$

where $\overline{u}=b\tilde{u}$ and $E(\overline{u})=0$, $Var(\overline{u})=b^2\sigma^2=\tilde{\sigma}^2$.

By estimating this equation, we can derive the underlying structural parameters: given c_0, c_1, c_2, c_3, we can estimate *b*, thus obtaining estimates also of ρ, β and *a*. From the latter parameter, whenever the value of $\frac{(\zeta_E-g_E)}{(\zeta_A-g_A)}$ is estimated or fixed in other model equations[22], we can also compute δ.

4.1.3. Capital Accumulation

Although equation (12e) is directly derived from a simple accounting identity, we still need some assumptions to express it in an estimable, i.e. linear, form. Firstly, we can explicitly express (12e) in the two-sector case as:

(12f)

$$\frac{k_t}{k_{t-1}}=\frac{(1-d)\overline{P}_{At}\,Y_{At}\left[(1+g_{At}-\varsigma_{At})+(1+g_{Et}-\zeta_{Et})\frac{(1-S_A)}{S_A}\right]R_{At}^{\alpha_2}\,e^{\phi t}u_t}{L_t\,k_{t-1}}+\frac{1}{(1+n_t)}$$

Secondly, we can reasonably assume that both the left hand side and the second term of the right hand side of equation (12f), though varying over time, tend to be quite regular in the long-run, that is $E\left(\frac{k_t}{k_{t-1}}\right)=h_1$ and $E\left(\frac{1}{1+n_t}\right)=h_2$ $\forall t=1,...,T$, where h_1 and h_2 are constant terms.

It follows that:

22 In particular, by assuming that $(\zeta_A-g_A)=1$, as explained below, we have $\frac{(\zeta_E-g_E)}{(\zeta_A-g_A)}=(\zeta_E-g_E)$ and this latter term can be estimated in (12h).

$$E\left(\frac{(1-d)\overline{P}_{At}\,Y_{At}\left[(1+g_{At}-\zeta_{At})+(1+g_{Et}-\zeta_{Et})\frac{(1-S_A)}{S_A}\right]R_{At}^{\alpha_2}\,e^{\phi t}u_t}{L_t\,k_{t-1}}\right)=h_3,\ \forall t=1,...,T,$$

where h_3 is a constant term, too. As it is always possible to find a value c for which, in general, $a+b=ac$ (where a, b and c are generic constant terms), we can adopt the following approximation:

(12g) $E\left(\frac{k_t}{k_{t-1}}\right)=\Delta+h_2\cong\Delta c\,e^{\varepsilon_{it}}, \quad \forall i=1,...,N$

where $\Delta=E\left(\frac{(1-d)\overline{P}_{At}\,Y_{At}\left[(1+g_{At}-\zeta_{At})+(1+g_{Et}-\zeta_{Et})\frac{(1-S_A)}{S_A}\right]R_{At}^{\alpha_2}\,e^{\phi t}u_t}{L_t\,k_{t-1}}\right)$, c is an unknown constant term, and ε_{it} is an error term with $E(\varepsilon_{it})=0,\ Var(\varepsilon_{it})=\sigma_\varepsilon^2, \forall i=1,...,N$ and $\forall t=1,...,T$.

If we replicate the approximation of $\frac{(1-S_A)}{S_A}$ described in Appendix 2, and assume that the agricultural sector does neither produce intermediate goods (i.e., $g_{At}=0$) nor shows negative or positive net export (i.e., $\zeta_{At}=1$),[23] we can finally express equation (12g) in logarithmic form as follows:

(12h) $\ln k_t=\tilde{a}+\ln\overline{P}_{At}+\ln y_{At}+c_1S_{At}+c_2S_{At}^2+c_3S_{At}^3+\alpha_2\ln R_{At}+\phi t+e_t$

where $\tilde{a}=[\ln(1-d)+\ln c+\ln(1+g_{Et}-\zeta_{Et})]$, c_0,c_1,c_2,c_3 are known parameters, $y_{At}=\frac{Y_{At}}{L_t}$ and $e_t=[\ln u_t+\varepsilon_t]$ is the disturbance term with $E(e_t)=0$ and $Var(e_t)=\sigma_\varepsilon{}^2+\varpi_u^2$ by assuming that $Cov(\varepsilon_t,\ln u_t)=0$.[24]

---

[23] The first assumption is largely acceptable, as most agricultural products are indeed aimed at final consumption while feed products, that are indeed intermediate goods, are consumed by agriculture itself and not by the non-agricultural sector. The second assumption is more problematic as some regions may have more export propensity than others. However, we may still think that this different propensity does exist, therefore $\zeta_{At}\neq 1$, but this non-zero net export is a stochastic term eventually included in the error term of (12h).

[24] We can also notice, that (12h) could be written as follows:

$$\ln k_t-\ln y_{At}=\ln\frac{k_t}{y_{At}}=\ln\frac{K_t}{Y_{At}}=\tilde{a}_i+\ln\overline{P}_{At}+c_1S_{At}+c_2S_{At}^2+c_3S_{At}^3+\alpha_2\ln R_{At}+\phi t+e_{it}.$$

However, to maintain as dependent variable the same variable appearing on the right hand side of (15c), we prefer specification (12h).

4.2. Region-specific Terms: the Panel-data Specification

From the derivations above, we conclude that the empirical specification of the model consists of a system of 3 simultaneous equations, (15c), (17f) and (12h), where only R_{At} and y_{At} are exogenous variables.[25] One further possibility to define the proper model specification, however, concerns the inclusion of region-specific (individual) terms. Introducing time and regional indices, *t=1,...,T* and *i=1,...N*, respectively, we can simply add region-specific intercepts (or Fixed Effects, FE) in three equations as follows:

(15d) $$S_{Ait} = \alpha_{Ai} - \alpha_{AE} \ln \overline{P}_{Ait} + \alpha_{AE}\, t\, (\lambda_E - \lambda_A) - \alpha_{AE}\alpha_1 \ln R_{Ait} + \delta_{AK} \ln k_{it}$$

(17g) $$\ln \overline{P}_{Ait} = (a_i + bc_0) + bc_1 S_{Ait} + bc_2 S_{Ait}^2 + bc_3 S_{Ait}^3 - b\, t \ln \beta + \overline{u}_{it}$$

(12i) $$\ln k_{it} = \widetilde{a}_i + \ln \overline{P}_{Ait} + \ln y_{Ait} + c_1 S_{Ait} + c_2 S_{Ait}^2 + c_3 S_{Ait}^3 + \alpha_2 \ln R_{Ait} + \phi t + e_{it}$$

In (15d), we assume that term $(\lambda_E - \lambda_A)$ is constant over time and regions, thus behaves as an unknown parameter. The fixed region-specific term α_{Ai} implies a region-specific revenue function and, consequently, region-specific PPF and technology.

In (17g), the individual term $a_i = \dfrac{\ln\left(\dfrac{\delta_i}{1-\delta_i}\right) + (1+\rho)\ln(\zeta_E - g_E)}{(2+\rho)}$, due to the individual distribution parameter δ_i, defines a region-specific utility function, thus, specific consumers' preferences and income elasticity.[26] Moreover, the error term itself, $\overline{u}_i = b\widetilde{u}_i$, with $E(\overline{u}_{it}) = 0$, $Var(\overline{u}_{it}) = b^2\sigma_i^2 = \widetilde{\sigma}_i^2$, becomes region-specific (see Appendix 2 for details).

Finally, in (12i) the region-specific term is $\widetilde{a}_i = [\ln(1-\delta) + \ln c_i + \ln(1 + g_{Et} - \zeta_{Et})]$, while $e_{it} = [\ln u_t + \varepsilon_{it}]$ is the disturbance term, with $E(e_{it}) = 0$ and $Var(e_{it}) = \sigma_{\varepsilon i}{}^2 + \overline{\sigma}_{ui}^2$. The presence of term $\ln c_i$ makes capital accumulation itself region-specific. This term comes from the

---

[25] Actually, we can also write $y_{At} = \dfrac{Y_{At}}{L_t} = S_{At}\dfrac{Y_t}{L_t} = S_{At} y_t$, where only variable y_t (the overall labour productivity) is exogenous. We could thus express (12h) as follows: $\ln k_t = a_i + \ln \overline{P}_{At} + \ln y_t + \ln S_{At} + c_1 S_{At} + c_2 S_{At}^2 + c_3 S_{At}^3 + \alpha_2 \ln R_{At} + \phi t + e_{it}$. This specification, however, becomes less empirically tractable for the presence of S_A in the logarithmic form. Moreover, both terms Y_t and Y_{At} are endogenous in the present model and, eventually, both y_t and y_{At} may be assumed exogenous only for the exogeneity of L_t. Thus, for simplicity, we prefer to directly consider y_{At} exogenous.

[26] Implicitly in (17g) we assume that $\dfrac{Y_{Ai}}{(Y_{Ai} - \gamma_A)} \cong (\beta)^{1/t}$ and $\dfrac{Y_{A0i}}{(Y_{A0i} - \gamma_A)} = (\beta)$, $\forall i=1,...N$.

approximation in equation (12g), so we can rewrite it as $E\left(\frac{k_{it}}{k_{it-1}}\right)=\Delta+h_{2i}\cong\Delta c_i\, e^{\varepsilon_{it}}$, thus admitting a region-specific demographic growth: $E\left(\frac{1}{1+n_{it}}\right)=h_{i2}$. It follows that, within the present model, the region-specific component of capital accumulation actually comes from the specific labour-force growth rate.

To complete the panel specification, we might also include a fixed region-specific effect in slope coefficients. This seems particularly interesting with respect to public agricultural R&D, R_{Ait}. This variable, in fact, is problematic for two major reasons. Firstly, we only have data about the aggregate public agricultural R&D expenditure observed at the national level, R_{At}. It includes both the regional and country-level research activities, and the latter can either deliver their effect homogeneously across regions or be strongly region-specific. Secondly, even if we had statistical information about the region-by-region R&D expenditure, nonetheless this would not correspond to the actual R&D that any region can exploit, as research done in one region can (and usually does) spill over other regions, especially the closer ones in geographical and structural terms.

We can still try, however, to divide R_{At}[27] in two components. The first concerns the region-specific expenditure, thus corresponding to N different shares on total expenditure; the second component applies to all regions homogeneously. We can thus write:

$$(18)\ R_{At}=\sum_{i=1}^{N}\omega_i R_{At}+\psi\, R_{At}\,,\quad \forall i=1,\ldots,N$$

where ω_i parameters indicate the i-th region-specific R&D component, while ψ indicates the nation-wide R&D component. Evidently, the following relation must hold: $\sum_{i=1}^{N}\omega_i+\psi=1$; it must be also true that $\omega_i\geq 0\ \ \forall i$, and $\psi\geq 0$. Then, from equation (18) we derive the expression for the region-specific public agricultural R&D, R_{Ait}:

$$(19)\ R_{Ait}=\omega_i R_{At}+\psi\, R_{At}=\left(\omega_i+\psi\right) R_{At}$$

Therefore, in equations (15d)-(12i) we actually include $\ln R_{Ait}=\ln\widetilde{\alpha}_i+\ln R_{At}$, where $\widetilde{\alpha}_i=\left(\omega_i+\psi\right)$ is a region-specific intercept term; as $\sum_{i=1}^{N}\omega_i=1-\psi$, we can also write

[27] Actually, in the present model, R_{At} represents the R&D stock which is calculated from yearly expenditure as described in Appendix 3.

$\sum_{i=1}^{N} \widetilde{\alpha}_i = 1 + (N-1)\psi$. According to this specification, the model now admits specific agricultural technological change due to a different slope associated to the national R&D. This, in turn, should affect both capital accumulation (negatively) and PPF (positively) that now assume a further region-specific component.

The $0 \leq \psi \leq 1$ parameter has an important economic interpretation as its magnitude actually indicates how increasing returns to R&D are. ψ being that public R&D expenditure replicating its effects over the N regions, it indicates how large is the non-rival part of national public R&D ($\sum_{i=1}^{N} \omega_i$ being the rival part), thus inducing increasing returns to scale at the aggregate level (Romer, 1986 and 1990; Barro, 1991; Fulginiti and Onofri, 2006 p.4).

4.3. The Dynamic Model

Equation (15d)-(12i) actually represent supply and demand side equilibria for goods A and E. Though we are willing to assume they strictly hold in the long-run, however, we should also admit that adjustment to new equilibria is not instantaneous and consequently some deviations from such equilibria can be observed in the short-run. In fact, if the reallocation of factors among sectors and of income across consumption goods in response to changes in technology, prices and aggregate factor endowments, is costly and occurs with some lag, the observed sectoral shares, relative price and aggregate capital accumulation, will reflect deviations from equilibria depicted in model equations (Coxhead and Punyasavatsut, 2002, p. 11). To realistically take into account these possible short-run deviations from long-run equilibria, we can specify a stochastic dynamic equilibrium model (Gemmell et al., 2000; p. 354).

A general dynamic specification of the simultaneous system of equations[28] can be formulated following Wickens and Breusch (1988; pp. 194-196). Their representation is particularly helpful as it emphasizes how many different forms this kind of models can actually take. The dynamic model can be written in compact matricial form as follows:[29]

$$(20)\ \sum_{s=0}^{T} \mathbf{A}_s \mathbf{Y}_{t-s} = \sum_{s=0}^{Z} \mathbf{B}_s \mathbf{X}_{t-s} + \mathbf{E}_t$$

28 This kind of models are also known as Dynamic Structural Equation Models (DSEM) (Lütkepohl, 1991; Zellner and Palm, 1974, 2004; Greene, 1997; Pagan and Robertson, 1998). Zellner and Palm (1974; 2004) also refer to these models as SEMTSA (Structural Econometric Time Series Analysis).

29 For simplicity, the presentation below will drop the panel-data notation (i.e., the i-th index).

where $\mathbf{Y}_t$ is the p×1 vector of endogenous variables, $\mathbf{X}_t$ is the q×1 vector of exogenous variables and $\mathbf{E}_t$ is the p×1 vector of disturbances assumed i.i.d. ($\mathbf{0}$,$\mathbf{\Omega}$). $\mathbf{A}_s$ and $\mathbf{B}_s$ are p×p and p×q matrices of unknown parameters, respectively. Parameter matrices differ over time index *s*, where $s \in 1,....,T$ or Z, and maximum lags T and Z may differ.

The reduced form of (20) can be easily derived:

$$(21)\ \mathbf{Y}_t = \sum_{s=1}^{T} \mathbf{C}_s \mathbf{Y}_{t-s} + \sum_{s=0}^{Z} \mathbf{D}_s \mathbf{X}_{t-s} + \mathbf{V}_t$$

where $\mathbf{C}_s = -\mathbf{A}_0^{-1}\mathbf{A}_s$, $\mathbf{D}_s = \mathbf{A}_0^{-1}\mathbf{B}_s$ and $\mathbf{V}_t = \mathbf{A}_0^{-1}\mathbf{E}_t$, that is, a vector of disturbances assumed i.i.d. ($\mathbf{0}$,$\mathbf{\Sigma}$). Under restriction $\mathbf{D}_s = 0\ \forall s > 0$, this specification corresponds to an unrestricted VAR model with exogenous variables,[30] while if $\mathbf{D}_s \neq 0$ at least for some *s*, (21) becomes a Partial Adjustment Model.[31]

Structural parameters in $\mathbf{A}_s$, $\mathbf{A}_0$ and $\mathbf{B}_s$ can not be identified from model (21) unless further identifying restrictions are imposed (see next section). Nonetheless, by appropriate transformations, it is possible to identify the matrix of the long-run multipliers, representing the structural effect of exogenous variables on endogenous variables. For this purpose, equation (21) can be reformulated as:

$$(22)\ \mathbf{Y}_t = \sum_{s=1}^{T} \mathbf{F}_s \Delta_s \mathbf{Y}_t + \mathbf{\Phi X}_t + \sum_{s=1}^{Z} \mathbf{D}_s \Delta_s \mathbf{X}_t + \mathbf{U}_t$$

where $\mathbf{F}_s = \mathbf{H C}_s$, $\mathbf{G}_s = \mathbf{H D}_s$, $\mathbf{U}_t = \mathbf{H V}_t$ and $\mathbf{H} = \left(\mathbf{I} - \sum_{s=1}^{T} \mathbf{C}_s\right)^{-1}$. The matrix of long run multipliers $\mathbf{\Phi}$ is calculated as: $\mathbf{\Phi} = \mathbf{H}\left(\sum_{s=0}^{Z} \mathbf{D}_s\right) = \left(\sum_{s=0}^{T} \mathbf{A}_s\right)^{-1}\left(\sum_{s=0}^{Z} \mathbf{B}_s\right)$. For simplicity we refer to (22) as the *Wickens-Breusch specification*.[32]

---

30 These models are also known as VARX models (Lütkepohl, 1991).

31 In fact, in the present application, we only consider the current value of exogenous variables, i.e. $\mathbf{X}_t$, thus with $\mathbf{D}_s = 0\ \forall s > 0$, even because the vector of exogenous variable actually contains predermined (time-invariant) or trend variables, thus non-stochastic terms.

32 Martin and Warr (1993, 1994) follow a long-run multipliers approach according to the abovementioned Wickens-Breusch specification, This specification is related to the so-called final form of the DSEM (Lütkepohl, 1991; Zellner and Palm, 1974, 2004; Greene, 1997; Pagan and Robertson, 1998) as both consider the whole adjustment process occurring in the dynamic model. In addition, by imposing restrictions on (22), for instance *T*=*Z* and $\sum_{s=1}^{T} \mathbf{C}_s + \sum_{s=0}^{T} \mathbf{D}_s = \mathbf{I}$, we can also give it a VECM interpretation, at least in the Hendry et al. (1984) specification. This representation is one of the possible ways to distinguish the long-run and short-run dynamics and may hold for both stationary and non-stationary and cointegrated series, the latter mostly being the cases for which the VECM is normally adopted. For instance, in presence of non-stationary and cointegrated series,

Here, though we admit short-run adjustments through the dynamic specification, we are mainly interested in the theoretical (i.e., structural) relations among the endogenous variables. Therefore, instead of adopting one of the mentioned reduced form specifications, we can reformulate the structural model (20) by isolating the contemporaneous effects as follows:

(23) $\mathbf{A}_0\mathbf{Y}_t = \sum_{s=1}^{T} - \mathbf{A}_s\mathbf{Y}_{t-s} + \sum_{s=0}^{Z}\mathbf{B}_s\mathbf{X}_{t-s} + \mathbf{E}_t$

Written in this dynamic form, model (15d)-(12i) can be interpreted as a Structural VAR (SVAR) model, where restrictions imposed on the VAR model (21) allow for the identification of the underlying structural parameters $\mathbf{A}_0$, $\mathbf{A}_s$ and $\mathbf{B}_s$. These identification issues will be discussed in detail in next sections.

4.4. The Estimated Model

By assuming a 2-year maximum lag,[33] we can now write (15d)-(12i) in the dynamic structural form (23) as follows:[34]

(24)

$$\begin{pmatrix} 1 & \alpha_{AE}^{s} & -\partial_{AK}^{s} \\ -b^{s}c_1 & 1 & 0 \\ c_1 & 1 & 1 \end{pmatrix}\begin{pmatrix} S_{Ait} \\ \ln \overline{P}_{Ait} \\ \ln k_{it} \end{pmatrix} = \begin{pmatrix} \rho_1 & \rho_2 & \rho_3 \\ \rho_4 & \rho_5 & 0 \\ \rho_6 & \rho_7 & \rho_8 \end{pmatrix}\begin{pmatrix} S_{Ait-1} \\ \ln \overline{P}_{Ait-1} \\ \ln k_{it-1} \end{pmatrix} + \begin{pmatrix} \rho_9 & \rho_{10} & \rho_{11} \\ \rho_{12} & \rho_{13} & 0 \\ \rho_{14} & \rho_{15} & \rho_{16} \end{pmatrix}\begin{pmatrix} S_{Ait-2} \\ \ln \overline{P}_{Ait-2} \\ \ln k_{it-2} \end{pmatrix}$$

$$+\begin{pmatrix} 0 & -\alpha_{AE}^{s}\alpha_1 & 0 \\ 0 & 0 & -b^{s} \\ 1 & \alpha_2 & 1 \end{pmatrix}\begin{pmatrix} y_{Ait} \\ R_{At} \\ c_2 S_{Ait}^2 + c_3 S_{Ait}^3 \end{pmatrix}$$

$$+\begin{pmatrix} \alpha_{AE}^{s}(\lambda_E - \lambda_A) & (\alpha_{Ai} - \alpha_{AE}^{s}\alpha_1 \ln\widetilde{\alpha}_i) & \tau_1 & \tau_2 \\ -b^{s}\ln\beta & (a_i + b^{s}c_0) & \tau_3 & \tau_4 \\ \phi & (\widetilde{a}_i + \alpha_2 \ln\widetilde{\alpha}_i) & 0 & 0 \end{pmatrix}\begin{pmatrix} t \\ D_i \\ CAP1_t \\ CAP2_t \end{pmatrix} + \begin{pmatrix} v_{it} \\ \overline{u}_{it} \\ e_{it} \end{pmatrix}$$

where $\mathbf{Y}_{it} = (S_{Ait}, \ln\overline{P}_{Ait}, \ln k_{it})$ is the (3x1) vector of endogenous variables; $\mathbf{X}_t = (\ln y_{Ait}, \ln R_{At}, (c_2 S_{Ait}^2 + c_3 S_{Ait}^3), t, D_i, CAP1, CAP2)$ is the (7x1) vector of exogenous

---

Coxhead and Punyasavatsut (2002) specify (22) as a VECM. However, this specification is not followed here as the adopted series are stationary (see Table 2).

33 In the empirical application, we actually admit lags varying from 0 (static model) to 3 years and then choose the optimal lag according to the usual procedures (see Table 2); eventually, all adopted specifications consider 2 lags. A 4-year lag or more would make the empirical specification hardly workable, so it is not considered here.

34 The panel-data notation (i.e., the i-th index) is now re-introduced, while, as explained in section 4.2., term $\ln R_{Ait}$ has been replaced by $\ln R_{At} + \ln\widetilde{\alpha}_i$.

variables,[35] here divided in two vectors, (3x1) and (4x1) respectively, to distinguish between stochastic time-varying exogenous variables, $\mathbf{X}_t^1 = \left(\ln y_{Ait}, \ln R_{At}, \left(c_2 S_{Ait}^2 + c_3 S_{Ait}^3\right)\right)$, and time-invariant or deterministic exogenous variables, $\mathbf{X}_t^2 = \left(t, D_i, CAP1, CAP2\right)$. D_i indicates the regional dummies taking account of the abovementioned regional fixed-effects. *CAP1* and *CAP1* are two time dummies expressing the introduction and the regime change of the Common Agricultural Policy (CAP), respectively (see section 6 and Appendix 3 for details); τ_1, τ_2, τ_3 and τ_4 are the respective parameters (hereafter synthetically indicated by the vector τ).[36] Finally, $\mathbf{E}_t = \begin{pmatrix} v_{it} \\ \bar{u}_{it} \\ e_{it} \end{pmatrix}$ is vector of multivariate structural disturbances assumed i.i.d. (**0**,**Ω**) and stationary over T.

With reference to equation (23), the matrices of structural parameters are:

$$\mathbf{A}_0 = \begin{pmatrix} 1 & \alpha_{AE}^s & -\partial_{AK}^s \\ -b^s c_1 & 1 & 0 \\ c_1 & 1 & 1 \end{pmatrix}, \quad -\mathbf{A}_1 = \begin{pmatrix} \rho_1 & \rho_2 & \rho_3 \\ \rho_4 & \rho_5 & 0 \\ \rho_6 & \rho_7 & \rho_8 \end{pmatrix}, \quad -\mathbf{A}_2 = \begin{pmatrix} \rho_9 & \rho_{10} & \rho_{11} \\ \rho_{12} & \rho_{13} & 0 \\ \rho_{14} & \rho_{15} & \rho_{16} \end{pmatrix},$$

$$\mathbf{B}_0^1 = \begin{pmatrix} 0 & -\alpha_{AE}^s \alpha_1 & 0 \\ 0 & 0 & -b^s \\ 1 & \alpha_2 & 1 \end{pmatrix}, \quad \mathbf{B}_0^2 = \begin{pmatrix} \alpha_{AE}^s \left(\lambda_E - \lambda_A\right) & \left(\alpha_{Ai} - \alpha_{AE}^s \alpha_1 \ln \tilde{\alpha}_i\right) & \tau_1 & \tau_2 \\ -b^s \ln \beta & \left(a_i + b^s c_0\right) & \tau_3 & \tau_4 \\ \phi & \left(\tilde{a}_i + \alpha_2 \ln \tilde{\alpha}_i\right) & 0 & 0 \end{pmatrix}.$$

As mentioned, c_0, c_1, c_2, c_3 are known parameters, while $\boldsymbol{\rho} = \left(\rho_1 \quad \rho_2 \quad \rho_3 \quad \rho_4 \quad \rho_5 \quad \rho_6 \quad \rho_7 \quad \rho_8 \quad \rho_9 \quad \rho_{10} \quad \rho_{11}\right)$ is the vector of the unrestricted parameters associated to lagged endogenous variables. It may be noticed that $\mathbf{A}_0$ is not lower triangular, that is, the system of equations is not recursive and this raises relevant issues for both the identification and estimation of $\mathbf{A}_0$, as discussed in next sections.

There is, in fact, a major problem in giving economic interpretation to this model specification. In (24) the long-term relations among the endogenous variables are a combination of parameters contained in $\mathbf{A}_0$, $\mathbf{A}_1$ and $\mathbf{A}_2$, that is of contemporaneous and

---

35 Though included in the vector of exogenous regressors, evidently S_{Ait}^2 and S_{Ait}^3 are indeed endogenous variables themselves, and this can create problems in the estimation of (24). Next section will discuss how to deal with this econometric issue.

36 The assumption here made about the role of the CAP is that it may affect both agricultural price and agricultural production, thus having a direct impact on both $\ln \overline{P}_{Ait}$ and S_{Ait} but not on the aggregate capital accumulation, $\ln k_{it}$. More detailed information about the regional CAP expenditure over time could be used instead of these simplistic dummies. However, available regional data only cover the period 1988-2000 (Esposti, 2007) and thus can not be adopted in the present study.

lagged effects. Consequently, coefficients in $\mathbf{A}_0$ do not correspond to the analogous coefficients in (15d)-(12i): $\alpha^s_{AE} \neq \alpha_{AE}$, $\partial^s_{AK} \neq \partial_{AK}$ and $b^s \neq b$. Though they express the same kinds of theoretical relation, the former parameters only capture the contemporaneous effects. To obtain the full long-run relations of (15d)-(12i), $\mathbf{A}_0$ parameters have to be integrated with the lagged effects (that is, the vector of parameters $\boldsymbol{\rho}$). For instance, α_{AE} in (15d) corresponds to $\alpha^s_{AE} + \rho_2 + \rho_{10}$ in (24). These longer-term effects can be eventually represented and assessed through the cumulated value of IRFs, upon stationartity of the series, though these actually combine all the complex interdependencies among endogenous variables (see section 7.3).

In practice, to identify these long-run relations, we have alternative solutions. On the one hand, if the modelled series are I(1) and cointegrated, we can adopt a VECM representation, where the cointegrating vector just identifies the long-run linkages among the endogenous variables. On the other hand, if series are stationary, the long-run relations can be just detected by estimating an unrestricted (reduced-form) VAR model, as (21), and computing the Impulse Response Functions (IRFs). The problem of imposing the theoretical restrictions in (24), however, remains critical in both cases.[37] If we think they should hold both in the simultaneous and longer-term effects, they may be imposed, or tested, in $\mathbf{A}_0$ as well as in the combination of $\mathbf{A}_0, \mathbf{A}_1$ and $\mathbf{A}_2$.[38] In any case, imposing these restrictions is eventually related to how we identify the SVAR model represented in (24).

5. Econometric Issues

5.1. Estimation

In this section we analyze in detail the econometric issues concerned by the estimation of (24) and the estimation strategy consequently adopted. These issues generally concern the contemporaneous presence of simultaneity (i.e., endogeneity), dynamics (i.e., AR terms) and regional heterogeneity (i.e., panel-data specification).[39]

---

37 As theoretical restrictions we should also consider symmetry and homogeneity imposed on technology and preferences. These restrictions, however, are already incorporated in (24).

38 For instance, in (24) $\ln \overline{P}_{Ait}$ does depend neither on $\ln k_{Ait}$ nor on $\ln k_{Ait-1}$ and $\ln k_{Ait-2}$.

39 Hsiao (2003) analyses the appropriate estimation approach of multiequational dynamic model (VAR, in particular) in the panel-data case. Details about the estimation of simultaneous equations using panel data, but with no specific reference to dynamic specifications, can be also found in Cornwell et al. (1992) and Baltagi (2005).

Actually, a larger heterogeneity across regions could be achieved by simply estimating the model region-by-region. On the one hand, however, the use of available information would not be particularly efficient in this case, as the number of parameters to be estimated would greatly increase even when they were not statistically different across regions. On the other hand, we prefer to approach the problem from the opposite point of view; that is, looking for common patterns (i.e., common parameters) still admitting some degree of regional heterogeneity, and eventually looking at differences and statistical significance among individual terms (see Appendix 4).

The estimation strategy here proposed is made of 4 steps and 4 different model specifications, according to an increasing degree of complexity in both estimation and identification issues. The first step tests for the presence of unit roots and, then, of cointegration among model variables. Among possible alternative unit root tests proposed for panel data (Baltagi, 2005; Hsiao, 2003), here non-stationarity is assessed through the IPS test (Im et al., 2003), widely used in the empirical literature and particularly appropriate when N is fixed and T tends to be greater (possibly, $T \rightarrow \infty$).[40]

The second step involves the estimation of model (24) in the reduced-form, that is a VAR model as (21) (or a VECM if cointegration is observed) (*model 1*). This reduced form model is estimated equation-by-equation by firstly using a conventional LSDV, or Within, estimator and, then, a GMM estimator (Arellano and Bond, 1991).[41] In fact, the usual VAR estimation (Lütkepohl, 1991) is not necessarily valid (namely, unbiased) under panel specification due to the so-called Nickell bias (Arellano, 2003, p. 85; Roodman, 2006, p. 13 and 31), the same problem occurring in univariate panel-data AR models (Esposti, 2007).

LSDV (or Within) estimate of an AR or VAR model is consistent whenever T goes to infinity (Arellano, 2003, p. 90), but it is biased in the small-sample and this bias may be large. In the present case, with N=20 and T = 52, bias is expected to be small (Vallanti, 2007, p. 15; Roodman, 2006, p. 13 and 31). Moreover, though GMM estimation should prevent this bias in principle, its small-sample performance is unpredictable and several practical aspects (namely, the choice of instruments) may be particularly critical in this respect (Hsiao, 2003;

---

40 A possible extension could be the adoption of the so-called second-generation tests that also admit cross-sectional dependence (Baltagi, 2005). Here, for simplicity, we prefer to maintain the hypothesis of cross-sectional independence.

41 We always refer to the so-called GMM-SYS estimator while the GMM-DIF estimate is not considered here (see Arellano, 2003, for more details on this aspect, and Esposti, 2007, for a recent application of both estimators).

Arellano, 2003, pag. 120; Holtz-Eakin et al., 1988).[42] Therefore, here we still perform both the LSDV (Within) and GMM estimations.

The equation-by-equation estimation (either with LSDV or GMM), however, can be inefficient compared to system estimation. Thus, we can also alternatively adopt an iterated-SUR estimation (with dummy variables) (Baltagi, 2005), or a joint-GMM estimation made on the three equations stacked together (Hsiao, 2003; Arellano, 2003, p.116-120). It must be reminded, however, that even when error term correlation occurs across equations, these system estimates should not differ from the equation-by-equation estimates unless equations include different regressors (Greene, 1997, pp. 488-489; Veerbek, 2004, p. 322). Consequently, it is useful to apply a system estimator only under exclusion restrictions, that is, when a restricted VAR is specified. Here, we adopt this solution on a VAR model where all restrictions specified in (24) are imposed accordingly on the lagged variables (*model 2*).[43]

The third step moves towards the identification of structural and long-term relations, though only with respect to exogenous variables and without imposing all the theoretical restrictions in (24), but only those referring to exogenous variables. It consists of the estimation of the Wickens-Breusch specification (*model 3*) adopting the appropriate panel-data system estimate (Cornwell et al., 1992; Baltagi, 2005; Hsiao, 2003; Wickens and Breusch, 1988). In this case, the lagged endogenous variables enter the right hand side of the equations as first differences; both $\Delta_1 \mathbf{Y}_t$ (that is $(\mathbf{Y}_t - \mathbf{Y}_{t-1})$) and $\Delta_2 \mathbf{Y}_t$ (that is $(\mathbf{Y}_t - \mathbf{Y}_{t-2})$; the latter can be also written as $(\mathbf{Y}_t - \mathbf{Y}_{t-1})$-$(\mathbf{Y}_{t-1} - \mathbf{Y}_{t-2})$) are correlated with the error terms in panel-data specification, and both are endogenous as they contain current values of the endogenous variables. Consequently, upon selection of appropriate instruments, we perform a 3SLS estimation (Martin and Warr, 1993; 1994). In this respect, it must be reminded that under a mean stationary AR(2) process, the lagged differences Δy_{it-s}, with s>2, are uncorrelated with ε_{it} (Blundell and Bond, 1998; Esposti, 2007). Thus, we can use Δy_{it-s}, with s $\geq$3, as legitimate instruments for $\Delta_1 \mathbf{Y}_t$ and $\Delta_2 \mathbf{Y}_t$.

---

42 For completeness, we report both the one-step and the two-step GMM estimators (Arellano, 2003). In fact, though asymptotically efficient, the two-step GMM estimates actually show significantly downward biased standard errors in small samples. Therefore, particularly for statistical inference, an appropriate correction should be made accordingly (Windmeijer, 2005).

43 These restrictions on the endogenous variables, being imposed on the reduced-form (that is, on the lagged variables), do not evidently correspond to the structural form restrictions indicated in (24) and that will be imposed in SVAR model estimation. In addition, in this restricted specification, variables with non-significant parameters in the unrestricted VAR could also be dropped. In this case, this restricted model is also called Parsimonious VAR (Hendry, 1995).

The final step concerns the estimation of the simultaneous equation model in its structural form (SEM), thus admitting contemporaneous effects among endogenous variables and imposing theoretical restrictions on structural parameters according to (24) (*model 4*). In doing this, firstly we have to deal with regressors S_{At}^2 and S_{At}^3. These variables are indeed endogenous and they may also raise problems due to the non-linearity. Thus, S_{At}^2 and S_{At}^3 are instrumented using the lagged values S_{At-1}^2 and S_{At-1}^3 as they are exogenous in the SVAR model ($E\left(S_{At-1}^2\ \varepsilon_t\ \right)=0$ and $E\left(S_{At-1}^3\ \varepsilon_t\ \right)=0$).[44]

Secondly, estimating a SEM within a dynamic panel-data specification raises several relevant econometric issues whatever approach is followed (namely, 3SLS, FIML or GMM) (Baltagi, 2005; Hsiao, 2003; Wickens and Breusch, 1988). A feasible alternative consists in estimating (24) as a SVAR model: first of all, the corresponding VAR model, that is (21), is estimated, and then, under appropriate identifying restrictions, structural parameters are obtained. To achieve this, however, relevant identification problems have to be solved.

5.2. Identification

Both VAR (unrestricted or restricted) and Wickens-Breusch specifications actually exclude the simultaneous relations, thus eliminating the identification issue by definition. Most structural parameters of (24), however, cannot be derived directly from the estimation of these reduced forms.

In this respect, we can distinguish structural parameters in two major groups. On the one hand, parameters that are explicit in (24) (for simplicity, we call them explicit parameters) and are: α_{AE}^s, δ_{AK}^s, b^s, α_2, ϕ, $\boldsymbol{\rho}$ and τ. On the other hand, parameters incorporated in other model parameters and that can be computed only indirectly (implicit parameters): α_1, $\left(\lambda_E-\lambda_A\right)$, α_{Ai}, $\ln\widetilde{\alpha}_i$, $\ln\beta$, a_i and $\widetilde{a}_i$. In particular, from the estimated region-specific intercepts of (24), $(\alpha_{Ai}-\alpha_{AE}\alpha_1\ln\widetilde{\alpha}_i)$, $(a_i+b^s c_0)$ and $\left(\widetilde{a}_i+\alpha_2\ln\widetilde{\alpha}_i\right)$, given that $\widetilde{a}_i=\left[\ln(1-d)+\ln c_i+\ln\left(1+g_{Et}-\zeta_{Et}\right)\right]$ and $a_i=\dfrac{\ln\left(\dfrac{\delta_i}{1-\delta_i}\right)+(1+\rho)\ln\left(\zeta_E-g_E\right)}{(2+\rho)}$, we can also compute the value of d, c_i, δ_i, ρ (and σ), ζ_E-g_E. This can be achieved reminding that $\widetilde{\alpha}_i=\left(\omega_i+\psi\right)$; these two parameters have a specific economic interpretation in terms of

---

[44] Greene (1997; p. 733 and p. 750) provides some suggestion on how to proceed in this case; other details can be found in Strickland and Weiss (1976).

rivality of the national public agricultural R&D (R_{At}) among regions, and assumptions can be thus made about their value.[45]

Identification of a SEM is a classical econometric issue (Greene, 1997). Eventually, we need enough restrictions, particularly on the simultaneous terms, to achieve the identification of all structural parameters. Interpreting (24) as a SVAR model, however, raises identification issues from a different perspective, the main problem being, following (21) and (23), to find a way to orthogonalize the reduced form error terms $\mathbf{V}_t$ ~($\mathbf{0}$, Σ) such that, $\mathbf{A}_0\mathbf{V}_t = \mathbf{L}\overline{\mathbf{E}}_t$, where $\overline{\mathbf{E}}_t$ are the orthogonal error terms but with $\overline{\mathbf{E}}_t$ ~($\mathbf{0}$, $\mathbf{I}$), and $\mathbf{L}$ a non-singular diagonal matrix of unknown parameters. $\mathbf{L}$ rescales the orthogonalized (structural) disturbances such that $\mathbf{E}_t = \mathbf{L}\overline{\mathbf{E}}_t$, and the consequent IRFs to this structural innovations can be interpreted as response to unit increases.

Pagan and Robertson (1998), Lastrapes and McMillin (2004) and Lucchetti (2006) present excellent descriptions of these identification issues in SVAR model estimation. On the one hand, a largely adopted solution relies on the Choleski factorization, namely a short-run restriction imposing recursiveness, that is, $\mathbf{A}_0$ to be lower triangular.[46] Alternatively, another kind of identifying restrictions still concerns the contemporaneous (short-run) effects and makes use of equation $\mathbf{\Sigma} = \mathbf{A}_0^{-1}\mathbf{L}\left(\mathbf{A}_0^{-1}\mathbf{L}\right)'$ implied by $\mathbf{A}_0\mathbf{V}_t = \mathbf{L}\overline{\mathbf{E}}_t$; putting enough restrictions on parameters of both matrices $\mathbf{A}_0$ (therefore, on the contemporaneous effects among endogenous variables) and $\mathbf{L}$, the remaining free parameters can be thus estimated (this is the so called *AB model*).[47]

On the other hand, we can impose long-run identifying restrictions (i.e., also involving the lagged-term parameters) through matrix $\mathbf{C} = \overline{\mathbf{C}}^{-1}\mathbf{L}$, where $\overline{\mathbf{C}}$ is the matrix of parameters of reduced-form cumulated effects, namely, following (21), $\sum_{s=0}^{T}\mathbf{C}_s = \sum_{s=0}^{T} - \mathbf{A}_0^{-1}\mathbf{A}_s$. By imposing restrictions on $\mathbf{C}$, therefore, we constraint the long-run cumulated effect of one endogenous variable on another. Putting enough appropriate restrictions and always basing on

45 The further assumptions required for the identification of the implicit parameters will be clarified in section 7.

46 The Choleski decomposition may be actually obtained according to the *n!* (*n*=3, in the present case) different Wold orderings. Thus, we can obtain *n!* different recursive VAR models and *n!* different sets of IRFs. This can make their interpretation difficult and controversial (Stock and Watson, 2001, p. 112).

47 In the literature matrix $\mathbf{L}$ is, in fact, normally named $\mathbf{B}$. Here, however, to avoid confusion with parameter matrices in (24), we do not use the conventional denomination $\mathbf{B}$.

$\mathbf{\Sigma}=\mathbf{A}_0^{-1}\mathbf{L}\left(\mathbf{A}_0^{-1}\mathbf{L}\right)'$, we can estimate the remaining free parameters in $\mathbf{C}$ (this is the so-called *C model*) (Pagan and Robertson, 1998; Lucchetti 2006).

In both cases, eventually, the SVAR identification implies a further estimation stage. After having estimated the reduced form (VAR), that is, $\hat{\mathbf{\Sigma}}$, straightforward restrictions must be imposed either on $\mathbf{A}_0$ and $\mathbf{L}$ or on $\mathbf{C}$; namely, exclusion restrictions or simple linear restrictions on the short-run (contemporaneous) or long-run (contemporaneous and lagged) effects.[48] Then, by exploiting relation $\mathbf{\Sigma}=\mathbf{A}_0^{-1}\mathbf{L}\left(\mathbf{A}_0^{-1}\mathbf{L}\right)'$, we can construct the respective log-likelihood functions, L($\mathbf{A}_0$, $\mathbf{L}$) and L($\mathbf{C}$), depending on $\mathbf{\Sigma}$ and on the parameters of $\mathbf{A}_0$ and $\mathbf{L}$ or $\mathbf{C}$. If restrictions imposed on these parameters are enough and appropriate (that is, we have an identified model), we can estimate the remaining free parameters by maximizing the respective log-likelihood function, i.e. by MLE (Amisano and Giannini, 1997).[49]

In the present application, between the two models (AB and C), we prefer to adopt short-run identification, although theoretical restrictions in (24) would allow both specifications. In fact, we are mainly interested in long-run relations, thus we prefer to let them free and eventually test them upon model estimation.[50] Therefore, the identification of the SVAR is achieved by relying only on short-run (contemporaneous effects) restrictions eventually allowing for (but not imposing) the validity of the long-run relations implied by the theory. It follows that we obtain structural parameter estimates by firstly estimating the restricted VAR model;[51] secondly, we impose restrictions on the contemporaneous effects (i.e., $\mathbf{A}_0$) as implied by (24)[52] and, then, the free parameters of $\mathbf{A}_0$ are estimated through MLE.[53]

---

48 Normally, it is not possible to jointly impose short and long-run restrictions.

49 ML estimation actually implies the assumption of normal distribution for $\mathbf{V}_t$, $\mathbf{E}_t$ and $\overline{\mathbf{E}}_t$.

50 Nevertheless, the SVAR model estimate obtained through the long-run identifying restrictions implied by (24) are available upon request.

51 This restricted VAR is slightly different from the previous SUR estimate (*model 2*), as we only impose restrictions on exogenous variables, while we exclude restrictions on $\ln k_{Ait}$ in $\ln \overline{P}_{Ait}$ equation and on the parameter *b* applied to $\left(c_2 S_{Ait}^2 + c_3 S_{Ait}^3\right)$. Thus, these restrictions are just tested *ex post*. It must be mentioned, however, that imposing restriction on the structural relations between endogenous and exogenous variables is not possible both in the *AB* model and in the *C* model. So the adoption of a short-run identification, here, is not decisive in this respect.

52 As mentioned, structural parameters in $\mathbf{A}_0$ are set at the theoretical long-term values but they actually only capture the contemporaneous effects. Therefore, we test, *ex post*, if the contemporaneous effect plus the lagged effects sum up to a value that is statistically correspondent to the theoretical value imposed in $\mathbf{A}_0$.

Then, we can also estimate structural parameters in $\mathbf{A}_s$ and $\mathbf{B}_0$, as $\mathbf{C}_s = -\mathbf{A}_0^{-1}\mathbf{A}_s$ and $\mathbf{D}_0 = \mathbf{A}_0^{-1}\mathbf{B}_0$, and this allows for the explicit testing of the long-run relations among variables as combinations of contemporaneous and lagged effects. Finally, the implicit structural parameters can be indirectly computed as discussed above.[54] For these latter indirect parameter estimates, however, we need to compute standard errors to run statistical tests. Here, standard errors are obtained through bootstrapping. Within a VAR model, bootstrapping requires particular attention as endogenous variables show cross-correlation and autocorrelation, so we might draw wrong bootstrap samples if we do not correctly take into account these underlying stochastic processes. Consequently, we follow the approach recently described by Swensen (2006)[55].

Firstly, we obtain the estimated residuals of the (unrestricted) VAR; these should be i.i.d.. Then, we can draw M bootstrap samples of these residuals ($\mathbf{V}_t^B$, $\forall t \neq 0$) and of the exogenous variables y_{Ait} and R_{At}; all pre-determined variables, such as trend and dummy variables, are replicated as such. For y_{Ait} and R_{At}, however, the bootstrap samples are extracted on the first differences, that is Δy_{Ait} and ΔR_{At}, as they may be either non-stationary or trend-stationary. Thus, the bootstrap y_{Ait} and R_{At} levels are recursively calculated fixing the observed starting values y_{Ai0} and R_{A0} and, then, using the bootstrapped first differences.

Finally, we can recursively compute the bootstrap vectors of endogenous variables $\mathbf{Y}_t^B$, $\forall t \neq 0$, using the bootstrapped errors $\mathbf{V}_t^B$, the bootstrap exogenous and predetermined variables ($\mathbf{X}_t^B$) and the estimated VAR coefficients.[56] Thus obtained, $\mathbf{Y}_t^B$ and $\mathbf{X}_t^B$ can be used

53 To identify a structural model from an estimated VAR, $2n^2\text{-}n(n+1)/2$ and $n^2\text{-}n(n+1)/2$ restrictions are needed in the *AB* and *C* models respectively, with *n* indicating the number of endogenous variables, namely the number of equations (Enders, 1995). As in a conventional SEM, this order condition is necessary but not sufficient. Thus, a further check of local identification is required. Here, we follow the method proposed by Amisano and Giannini (1997). Imposing all restrictions on short-run structural parameters of (24) and performing this identification check, we conclude that the present SVAR model is identified.

54 It must be reminded that, if the GMM estimation is used, the region-specific fixed effects and all the relative structural information outlined above, are dropped due to first differentiation. However, these parameters can be still recovered from model estimation as $\hat{\mathbf{a}}_i = \hat{\mathbf{Y}}_{it} - \mathbf{X}_{it}\hat{\boldsymbol{\beta}}$ (Dhrymes, 2006, p. 9, Arellano, 2003, p. 119).

55 See also Lütkepohl (1991, p. 497; 2000, p. 4), Jouini (2006, p.7) and Omtzigt and Fachin (2002, p. 7). For more details on bootstrapping see, for instance, Davison and Hinkley (1997) and Chernick (1999).

56 Actually, we only use statistically significant parameters while fix at 0 all the others. As starting values we use real observations. A problem may arise with $\left(c_2 S_{Ait}^2 + c_3 S_{Ait}^3\right)$. However, as these terms are instrumented by the lagged values as mentioned, we can use the first real observations for the initial

to estimate the restricted VAR and SVAR models again; this procedure is repeated many times to obtain an estimate of standard errors of the implicit parameters.[57]

6. Data Sources and Some Descriptive Evidence

In the last decade, new and broad dataset on Italian regional economies have been released. Of major relevance, here, are the CRENoS regional database (Paci and Saba, 1997) and the AGREFIT database (Rizzi and Pierani, 2006). These datasets cover the whole post-war period and all Italian regions. Consequently, data include, over time and across space, a wide range of different development stages, from "developing to developed regions" (Abbott and McCalla, 2002).

Appendix 3 describes in detail definition, sources and construction of all model variables. The final dataset includes the 20 Italian regions observed over the 1951-2002 period (1040 observations), the 3 endogenous variables $\left(\mathbf{S}_{\mathbf{A}t}, \mathbf{ln}\overline{\mathbf{P}}_{\mathbf{A}t}, \mathbf{lnk}_{t}\right)$ and all the exogenous variables, that is, $\left(\mathbf{ln\,y}_{At}, \mathbf{ln\,R}_{At}, \left(c_2\mathbf{S}^2_{Ait} + c_3\mathbf{S}^3_{Ait}\right)\right)$, the time trend, 19 regional dummies and the two time dummies *CAP1* and *CAP2*.

Table 1 and Figures 2-6 report descriptive figures about endogenous and exogenous variables. They show the relevant structural change experienced by Italian economy and agriculture in the post-WWII decades and, at the same time, the significant and persistent differences among regions. This emerges by looking at either the aggregate or regional average CV (Coefficient of Variation) (Table 1), or at the two extreme cases (Figures 2-5), the region with the lowest average agricultural share on regional GDP (i.e. Valle d'Aosta region) and the region with the highest share (i.e., Molise region).

Regional patterns are indeed quite similar and where differences are initially (i.e., in 1951) observed, they reduce over time for most variables. This is particularly evident for $\overline{P}_{At}$ and k_t . The former always decreases over time, though the decline is stronger in the second half of the period; regional courses are very similar in this respect, and the CV itself progressively declines to almost 0 in 2002. This behaviour, however, should not surprise as there is no reason to expect divergent regional prices within the same country, while, on the contrary, price convergence could be the result of an increasing market integration among regions.

---

recursive calculations of the endogenous, and, then, use the new computed values of S_{Ait} to instrument $\left(c_2 S^2_{Ait} + c_3 S^3_{Ait}\right)$ in the following calculations of the endogenous variables.

57 Standard errors of bootstrap parameters are computed applying the formula suggested by Efron (1982) that also considers the bias of the bootstrap estimate with respect to the real estimate of the parameter under study.

Also capital intensity shows quite a regular pattern over the years, as expected. It always increases, though the highest growth rates are observed in the first half of the period. In this case, regional convergence towards the national average seems questionable when looking at the extreme cases (Figure 2). Nonetheless, the CV significantly declines and the gap between the extreme cases and the national average, at least in % terms, largely reduces.

There is no evidence of convergence, indeed, for sectoral variables, that is S_{At} and y_{At}. Still we observe, in both cases, a quite regular course over time, with agricultural share constantly declining, but at a slightly diminishing rate, and the *per capita* agricultural production showing a sharp increase in the first 15 years and, then, becoming almost stable in the following period. Interestingly, this latter case is the only variable for which the CV across regions increases over the whole period, thus we observe a slight divergence, which is even clearer looking at the extreme cases (Figure 5).

For agricultural share regional convergence can be questioned, too. Apparently (Figure 4), in all regions this share drops below 5%; however, the CV is almost constant as the regional differences, in % terms, remain significant. This can be interpreted by the fact that, although agriculture remarkably and constantly declines over years, the relevance of the sector within any regional economy is, somehow, specific. Together with the evidence concerning y_{At}, this leads to the general conclusion that, while moving along a common and, at least partially, convergent development pattern, Italian regions maintain over time some degree of agricultural specialization.

A final consideration can be made on the other major exogenous variable within the model, R_{At}. In this case, as largely discussed in section 4.2, regional observations are not available and we can only rely on the national aggregate value. It is still interesting to notice that R&D stock, expressed in real terms, strongly increases over time at an annual growth rate (+6,6% on average) which is higher than all other model variables; in particular, about 2% greater than the average annual growth rate of k_t. Therefore, at least apparently, the public effort in this direction is remarkable. At the same time, however, its course changed over time, as we can distinguish three distinct sub-periods in this respect. In the first two decades, R_{At} increased quite regularly at a low growth rate; then, it started growing quite rapidly in the '70s and '80s, but suddenly stopped and even slightly declined in the last decade.

[Table 1 here]

[Figures 2-6 here]

7. Results

7.1. Unit Root Tests

Table 2 reports the results of IPS unit root test on all model variables; as mentioned, among alternative tests assuming cross-section independence, this test seems appropriate due to the characteristics of the panel itself (T>N).[58]

With respect to previous similar studies (Martin and Warr, 1993, 1994; Gemmel et al., 2000; Punyasavatsut and Coxhead, 2002; Sun et al., 2007), a major difference emerges: with the exception of R_{At} , all series are stationary, i.e. I(0).[59] Consequently, neither cointegration occurs among variables nor the VECM (or VAR in first differences) specification is needed, as a VAR specification in the levels is fully appropriate.

This result, however, should not surprise. Previous studies actually refer to other countries and periods[60] and, above all, are all based on time series, not on panel data. As well known (Baltagi, 2005), the properties of unit root tests, mainly their power, significantly change moving from the conventional ADF tests applied to time series towards the panel-data unit root tests here adopted.

The specification of the test may be itself critical, as may depend on the maximum number of lags and on the presence of a deterministic trend. Even in this case, however, Table 2 shows uncontroversial results, as the rejection of the unit root is confirmed for all series under all cases (1 or 2 lags; with or without trend).[61] Only R_{At} behaves differently, but even in this case we may explain this evidence with the substantial difference in the way the test is performed.

R_{At} , in fact, has only a time series dimension, as we do not observe regional data for it. Non-stationarity is tested through conventional ADF tests indicating that, regardless of the specification, the presence of unit root can not be excluded. Though non-stationary, however, this is an exogenous variable so it remains legitimate to specify the model as a VAR in the

---

58 Alternative options are the LLC (Levin, Lin and Chu) and the HT (Harris and Tzavalis) tests (Baltagi, 2005). The former is, however, more restrictive than IPS; this should be thus preferred, in general terms. The latter is not appropriate here, as it assumes fixed T and N going to infinite (thus N>T). However, in the present case, results of the HT test (available on request) fully confirm the IPS test.

59 Tests are actually performed on model variables, therefore all expressed in logarithms except S_{At} (Table 2). This, however, does not affect the conclusion as the same time-series properties hold in levels and in logarithmic transformations.

60 They all deal with Far-East Asian countries (Malaysia, Thailandia, Indonesia, Taiwan) over the recent decades.

61 Evidently, the same results would be observed introducing further lags; therefore, we do not report tests for 3 or more lags.

levels. The presence of non-stationary explanatory variables should not create problems of spurious regression, even considering the inclusion of a trend among regressors.

[Table 2 here]

7.2. Comparison of the Alternative Specifications

7.2.1. Reduced-Form Models

The OLS-LSDV and the GMM estimates of *model 1* are reported in Tables 3 and 4, respectively. It is interesting to shortly comment a very basic specification of the model, that is the static case where the three endogenous variables are only regressed on the exogenous ones (first column of Table 3). Most parameters are indeed statistically significant and largely confirm the descriptive evidence discussed in previous section. Relative price is negatively affected by the trend and positively by the R&D stock; the introduction of the CAP, mainly based on price support, favoured the agricultural relative price, whereas the 1992 Reform (roughly based on the reduction of price support with the introduction of direct payments) acted in the opposite direction. Similar evidence concerns capital intensity, the major difference being, as expected, the positive linkage occurring with the *per capita* agricultural production. On the contrary, the agricultural share, which is of major interest here, shows a negative, and significant, correlation with all exogenous variables with the only exclusion of the 1992 CAP reform, whose effect on sectoral share turns out to be positive and relevant.

It might surprise that in all cases the trend variable is associated to a negative parameter even when, as in the case of $\ln k_t$, the dependent variable has a clear growing pattern. This can be partially explained by the possible collinearity eventually occurring between the trend variable itself and other two exogenous variables, namely $\ln R_{At}$ and $\ln y_{At}$. In general terms, however, we can hardly give a clear-cut economic interpretation to these results, as the static model evidently disregard the dynamic relations among endogenous variables, and all these effects are eventually mixed-up in the observed relations with the exogenous variables.

The second and third columns of Table 3 report the estimates of the unrestricted VAR model, in the pooled-data (no regional heterogeneity admitted) and panel specifications, respectively. It may be noticed that results are similar in the two cases, such that they can be discussed together. In fact, both the adjusted-R^2 and the AIC indicate that the panel specification is only slightly preferable for all the equations.[62] This should not surprise, as for

[62] Following these indicators the 2-lag specification has been adopted, as confirmed in table 7.

$\ln \overline{P}_{At}$ regional patterns are very close, whereas in the case of $\ln k_t$ and S_{At} differences may be not so large to justify significantly different parameters.

The relation among endogenous variables suggests that aggregate capital intensity is not significantly affected by agricultural relative price and sectoral share. At the same time, the other way round also holds true. Beside the own-lagged effects, which are always the largest, the only significant linkage among endogenous variables concerns $\ln \overline{P}_{At}$ and S_{At} , though the direction of this relation is somehow controversial. In particular, while agricultural share positively responds to a lagged increase of relative price, this latter is negatively affected by the lagged increase of agricultural share.

It must be also acknowledged that, when a 2-year lag is considered, the two lagged variables act in opposite directions almost reciprocally offsetting, particularly in the case of cross-variable effects. The only exception is S_{At} whose own 2-lag effect actually reinforces the previous one, thus suggesting that for this variable the short-term adjustments do not correct but, indeed, reinforce its own longer-term course.

Notwithstanding the introduction of lagged endogenous variables, exogenous ones continue to be significant, though their effects are now smaller. The time trend negatively conditions both $\ln \overline{P}_{At}$ and $\ln k_t$, while CAP dummies act in the opposite direction. The R&D stock is positively related to capital intensity, and this effect is the largest among exogenous variables, as well as to agricultural sectoral share, which is also remarkably favoured by an increase in *per capita* agricultural production; this can be interpreted as the effect of agricultural specialization within regional economies.

In principle, under panel specification, the LSDV (or Within) estimator could generate biased results, though its size depends on T; this should not occur when a GMM estimation (Arellano and Bond, 1991) is performed. However, these desirable characters of the GMM estimates only concern asymptotic properties and do not necessarily hold in the small sample. Table 4 reports the GMM estimates of the unrestricted VAR, so these results can be directly compared to Table 3 results.

This comparison suggests that the one-step GMM estimate does not differ very much from the LDV results, both in sign and size of parameters, the only difference being that $\ln \overline{P}_{At-2}$ now negatively and significantly affects $\ln k_t$. On the contrary, the two-step GMM estimation clearly provides poorer results, at least in terms of statistical significance and for two equations, i.e. $\ln \overline{P}_{At}$ and $\ln k_t$, while for S_{At} the two estimations largely correspond. LM and Sargan tests also confirm that the dynamic specification of the model, i.e. AR(2) processes, is the most appropriate, as well as the selection of the instruments in GMM estimation. Under an AR(2) specification, in fact, both 1st order and the 2nd order autocorrelation should be observed. Here, indeed, the former clearly holds true, while the latter does not occur in the case of equation $\ln k_t$. Anyway, these tests suggest that the adoption of a VAR(2) specification is appropriate.

System estimates are reported in Tables 5 and 6 (iterated-SUR and joint-GMM estimation, respectively). As mentioned, these results differ from the previous ones only whenever exclusion restrictions are imposed. Here we report the estimation of *model 2* mentioned above, where we impose all restrictions among variables (either endogenous or exogenous) indicated in the theoretical model. Even in this case, SUR estimates are closer to the one-step GMM rather than the two-step GMM results and, as far as significant parameters are concerned, the empirical evidence largely corresponds between the two. The Sargan tests suggest the appropriate choice of instruments, whereas LM tests, while supporting the AR(1) specification, provide less clear indication with respect to the AR(2) case adopted here.

Compared to the unrestricted model, most results are confirmed. The relation between agricultural relative price and sectoral share remains the same and of the same magnitude, whereas now capital intensity seems also dependent on the 2-year lagged value of both agricultural share and relative price: $\ln k_t$ is negatively affected by an higher share of the primary sector, while $\ln \overline{P}_{At}$ operates in the opposite direction. It remains true, however, that lagged capital intensity has not effect on the other two endogenous variables. Exogenous variables, when admitted, are statistical significant in most cases, though the R&D stock remains the major (positive) factor. When admitted and significant, the deterministic trend operates as a negative shifter, while both CAP time dummies tend to push agricultural relative price upward.

Being reduced-form models, the estimates of both the unrestricted and restricted VAR can not be easily interpreted as they mix-up short and lung-run relations among variables, and, above all, conceal the structural linkages, that is what happens to endogenous variables if one endogenous or exogenous variable varies. A clearer picture of these complex relations among the endogenous variables, however, can be achieved presenting VAR estimates through Granger-causality tests and, above all, IRFs (Table 8 and Figure 7, respectively).[63]

Granger-causality tests largely confirm and clarify further what discussed about VAR parameter estimates. While capital intensity neither Granger-causes the other two variables

---

63 "Because of the complicated dynamics in the VAR, these statistics are more informative than are the estimated VAR regression coefficients or R^2 statistics, which typically go unreported" (Stock and Watson, 2001, p.104)

nor is caused by them, for $\ln \overline{P}_{At}$ and S_{At} the reciprocal causation is evident. Though these tests are not, indeed, very much informative about the underlying economic relations, they still confirm expectations, that is, whenever all the relevant theoretical restrictions are not appropriately imposed, $\ln k_t$ behaves as an exogenous factor, whereas the two sectoral variables, relative price and share, are indeed strongly interdependent.

The IRFs are helpful in giving a clearer structural interpretation to this evidence. As mentioned in section 5, they are obtained by imposing short-run identification through Choleski decomposition, that is, by imposing a recursive structure. As this is always arbitrary to some extent, Figure 7 reports the IRFs of all the different possible decompositions, the six Wold orderings. Evidently, if IRFs strongly differ across orderings, this would imply strong dependence on how recursiveness is imposed; otherwise, we may consider them as robust results.[64]

The IRFs confirm that any endogenous variable mainly responds to its own shocks; this response, however, is always quite limited in size and temporary for both $\ln \overline{P}_{At}$ and S_{At} , while tends to be more persistent for $\ln k_t$. In particular, S_{At} response becomes lower than .01 just after the first year, while for $\ln \overline{P}_{At}$, though regularly declining, it always remains higher than .01 for 10 years. In case of $\ln k_t$, the response after 10 years remains almost at the same level of the initial response. This result would suggest, not surprisingly, that capital intensity is, among endogenous variables, a cumulative process, as any variation is likely to reinforce itself and tends to be maintained in the longer-term. This does not happen for the two sectoral variables, particularly for S_{At} , on which any external shock indeed generates only short-run adjustments.

Also cross-response to shocks confirm previous comments. A shock on both $\ln \overline{P}_{At}$ and S_{At} is practically irrelevant for capital intensity, thus suggesting that a greater impulse to agricultural growth does not seem to improve, through capital accumulation, aggregate economic development. On the contrary, shocks on these two sectoral endogenous variables have a reciprocal, though quite small, positive effect: as could be expected according to Engel's law, an increase of agricultural relative price generates a limited but persistent response of the agricultural sectoral share; the other way round also holds true but this effect is almost negligible and changes direction across different orderings.

---

64 As all endogenous variables are here expressed as logarithms or shares and given that IRFs always express responses to unit shocks, in the present case we can read IRFs directly as the percentage response to a 1% shock.

It may be noticed, in addition, that a shock on capital intensity produces a small but increasing effect on both $\ln \overline{P}_{At}$ and S_{At}, though this impact acts in opposite directions: an increase in $\ln k_t$ produces a reduction of S_{At}, while an increase in capital intensity fosters $\ln \overline{P}_{At}$. The former result would actually support the idea of a lower capital intensity in the primary sector.

In any case, IRFs do not change very much across the different Wold orderings. It implies that the structural relations among the endogenous variables identified through recursiveness are quite stable, regardless the way recursiveness is imposed. This occurs also because one of the three endogenous variables, $\ln k_t$, behaves as an almost-exogenous variable, while the other two show strongly reciprocal relations. Nonetheless, the economic interpretation of these results requires caution. The structural model obtained through recursiveness, in fact, is not the theoretical model in (24) as respective restrictions do not correspond to the theoretical restrictions (see next section).

As a matter of fact, we may notice that *model 1* should be preferred to restricted VAR (*model 2*).[65] Table 7 adopts two major selection criteria (Log Likelihood and AIC) to compare different reduced-form specifications. Among unrestricted specifications, the 2-lag case is preferred, although the pooled model performs better than the 1 and 3-lag specifications; the 2-lag case will be thus maintained hereafter. The restricted model should be evidently dropped in favour of the unrestricted one. Again, however, it must be underlined that restrictions on which this model is based are not the theoretical restrictions underlying model displayed in (15d)-(12i) or (24).

The last reduced-form specification here presented is the Wickens-Breusch specification (*model 3*). It aims at identifying the long-term (i.e, structural) impact of exogenous variables on endogenous variables (Martin and Warr, 1993 and 1994). To avoid endogeneity bias a 3SLS estimator is adopted.[66] The overall goodness of fit in this case is much lower than VAR specifications above (Table 9). This emerges clearly by looking at either the adjusted-R^2 or the limited number of statistically significant parameters, especially concerning the differential terms. Among the three equations, however, S_{At} apparently provides the better

---

65 Table 8 and Figure 7 refer to the unrestricted model; as seen, however, the unrestricted and the restricted VAR estimates do not differ very much, therefore Granger-causality tests as well as IRFs of the restricted VAR largely confirm what is here reported for the unrestricted case. These further results are available upon request.

66 As we impose on exogenous variables restrictions indicated in (24), a system estimation (3SLS) should be preferred in this case to equation-by-equation estimation (2SLS) (Martin and Warr, 1993 and 1994; Baltagi, 2005, Cornwell et al., 1992). For the selection of instruments we followed Martin and Warr (1993 and 1994), as also outlined in section 5.1, though here within a fixed-effect panel specification.

results, at least in terms of significant parameters. In any case, it is not straightforward to compare these results to previous VAR estimates, as the way they represent model dynamics substantially differs. We can still conclude, however, that *model 3* estimate does not provide a clear picture of the structural relations among endogenous variables.

On the contrary, interesting results are obtained with respect to the long-term multipliers of exogenous variables, as these parameters should not mix-up short and long-run responses and should be independent on the linkages among endogenous variables. It may be noticed that CAP time dummies, when admitted, are significant and positive with respect to $\ln \overline{P}_{At}$ and $\ln k_{At}$, while, quite surprisingly, in the share equation the impact is negative in one case (*CAP1*), and not significant in the other (*CAP2*). Thus, ultimately, the effect on S_{At} of the introduction of the CAP is negative, while its 1992 reform is negligible.

Of greater relevance, here, are the two exogenous variables eventually generating the technological gap effect in the explanation of agricultural decline; that is, *t*, expressing the autonomous and exogenous technical change rate affecting both sectors, and $\ln R_{At}$, namely the agricultural specific technical change induced by public R&D. The parameter associated to the deterministic trend *t* is statistically significant and negative for $\ln \overline{P}_{At}$ and $\ln k_t$, while positive but not significant for S_{At}. Therefore, this component of technical change rate is not a relevant factor in the explanation of agricultural decline, presumably because it acts in the same direction and intensity in both sectors.

The effect of $\ln R_{At}$ on S_{At}, conversely, is significant and positive and almost double than *t*, even though, and quite surprisingly, much less relevant than the impact of $\ln R_{At}$ on capital intensity. As this component only acts on sector A, the combination of these two parameters should be interpreted against the technological gap explanation of agricultural decline. On the contrary, there should be a contribution to a reduction of agricultural decline coming from this technology effect.

[Tables 3-9 here]

[Figure 7 here]

7.2.2. Structural Form Specification

None of the specifications and the respective estimates presented so far corresponds to the theoretical model (24). As mentioned in section 5.2, the identification of the structural parameters can be actually achieved by estimating the restricted reduced-form (VAR) corresponding to (24) and then estimating the structural parameters in $\mathbf{A}_0$. The former estimate, providing the reduced-form parameter matrices $\mathbf{C}_1, \mathbf{C}_2, \mathbf{D}_0$, as well as $\hat{\mathbf{\Sigma}}$, is here obtained by a system estimation (iterated-SUR) of the VAR under the exclusion restrictions that (24) imposes on exogenous variables (i.e., on $\mathbf{D}_0$), the assumption being that restrictions holding in the structural model remain valid even in the reduced form.[67] Then, short-run restrictions indicated in (24) are imposed on $\mathbf{A}_0$ and the remaining free parameters of $\mathbf{A}_0$ are estimated, given $\hat{\mathbf{\Sigma}}$, through MLE. Table 10 reports the results of these two estimation stages of the SVAR model (*model 4*).

Beside free parameters of $\mathbf{A}_0$, however, we also have to estimate the other explicit parameter in matrices $\mathbf{A}_1, \mathbf{A}_2, \mathbf{B}_0^1, \mathbf{B}_0^2$. This indirect estimation can be easily obtained from the following relations: $\mathbf{C}_1 = -\mathbf{A}_0^{-1}\mathbf{A}_1$, $\mathbf{C}_2 = -\mathbf{A}_0^{-1}\mathbf{A}_2$, $\mathbf{D}_0^1 = \mathbf{A}_0^{-1}\mathbf{B}_0^1$ and $\mathbf{D}_0^2 = \mathbf{A}_0^{-1}\mathbf{B}_0^2$. Finally, the implicit parameters of the model, according to the derivation presented in section 4, are derived by simple calculations also involving the estimated fixed-effects, that is the region-specific parameters associated to dummy variables D_i. This set of other explicit and implicit parameters are reported in Table 11.

In Table 10, matrices $\mathbf{C}_1, \mathbf{C}_2, \mathbf{D}_0$ do not differ very much from Table 5; thus, we do not comment these results in detail, even considering that we can not give them a direct structural interpretation. We may notice, however, that the different restrictions and the introduction of term $\left(c_2\mathbf{S}_{\mathbf{A}t-2}^2 + c_3\mathbf{S}_{\mathbf{A}t-2}^3\right)$, whose parameter is statistically significant, still affect results on some relevant aspect. In particular, while the S_{At} equation is almost coincident to previous estimate (*model 2*), some differences emerge in the other two equations. Relative price reacts to changes in agricultural share but only after 2 years, while now capital intensity is more dependent on the other two variables being directly and negatively affected by $\ln \overline{P}_{At-2}$ and indirectly, with a 4-year lag, by S_{At} .

Of major relevance is the MLE estimate of the three free parameters in matrix $\mathbf{A}_0$: α_{AE}^s, ∂_{AK}^s and b^s. They are all statistically significant and can be interpreted as structural parameters though only, as mentioned, as contemporaneous relations among model variables. The long-term relations can be obtained here as full effects summing up these contemporaneous parameters with the lagged structural ones included in $\mathbf{A}_1$ and $\mathbf{A}_2$.

---

67 All the other restrictions, including those of the structural model, can not be imposed ex ante in the reduced form estimate. Beside the identifying restriction imposed on $\mathbf{A}_0$, they can be only tested *ex post*. In particular, the two parameters fixed at 0 in matrices $\mathbf{A}_1$ and $\mathbf{A}_2$ of (24), are free and indicated as $\rho_{Pk(-1)}$ and $\rho_{Pk(-2)}$, respectively.

α^{s}_{AE} is the key-parameter of the share equation; its negative value confirms a direct proportionality between agricultural share and relative price, thus supporting one of the basic stylized facts underlying Engel's law: sectoral share and relative price move in the same direction.

∂^{s}_{AK} is another parameter of the share equation, but it refers to factor use and proportion across the two sectors; consequently, it expresses magnitude and direction of the Rybczynski effect. The negative value here obtained suggests a lower agricultural capital intensity and, therefore, that this effect becomes a relevant factor in explaining agricultural decline.

Finally, parameter b^{s} concerns the inverse demand equation and should express how relative price reacts to changing sectoral shares. A positive value would suggest that agricultural relative price increases as effect of an increase in agricultural share, which is, in turn, attributable to a change in consumption patterns in favour of agricultural products.[68]

These structural parameters, however, do not provide a complete picture of the relations among variables. On the one hand, they only refer to the contemporaneous (short-run) effects, and do not necessarily correspond to the long-run effects. On the other hand, many relevant structural parameters are indeed implicit in these results and have to be computed indirectly to fully interpret these relations. On the former aspect, it is of particular interest to compute the IRFs implied by this SVAR model. Figure 8 reports these structural IRFs.

These response functions are remarkably different from those obtained from recursiveness in Figure 7, and provide very useful information on how structural linkages among endogenous variables have to be interpreted. First of all, the variable of major interest here, S_{At}, responds to own and cross-equation shocks with the same order of magnitude. Response to own shock is not the most relevant anymore; in addition, this response is significantly positive only in the first years, then it becomes negative after the fifth year, suggesting that temporary increase of agricultural share tends to progressively counterbalance itself over time.

A more surprising result is that the impact on sectoral share of a relative price shock is almost negligible: slightly positive in the first years and then statistically equal to 0 in the longer term. On the contrary, the major impact on agricultural share, thus on sectoral decline, is caused by capital intensity. After a significant positive effect in the first year, the response

---

68 It must be noticed, however, that the contemporaneous response of relative price to change in sectoral share does not depend only on b^{s}, as it is: $\frac{\partial \ln \overline{P}_{Ait}}{\partial S_{Ait}} = b^{s}c_1 + 2b^{s}c_2 S_{Ait} + 3b^{s}c_3 S^{2}_{Ait}$. For $S_{Ait} = .1$, for instance, and given the values of c_1, c_2, c_3, it would be $\frac{\partial \ln \overline{P}_{Ait}}{\partial S_{Ait}} = 4.7$, thus confirming the positive response.

then becomes statistically negative and much larger than other responses. Eventually, agricultural decline seems prevalently driven, in the longer term, by the Rybczynski effect, namely an aggregate capital intensification, thus by economic growth itself.

The IRFs of $\ln k_t$ is also of major interest. It is confirmed that the response to own shock is positive, almost stable in the longer-term and always greater than .8. This evidently suggests the cumulative character of the capital accumulation process. Capital intensity response to shocks on $\ln \overline{P}_{At}$ is, on the contrary, negative but very small and, after few years, not statistically different from 0.[69] This does not happen for the agricultural sectoral share whose effect on capital intensification is significantly positive and remains at the initial levels also in the longer term. This evidence supports the idea that, during its decline, which is itself fostered by capital intensification, agriculture gives a positive and persistent contribution to economic growth, though we can hardly say that it causes growth (at least, in the sense of Granger-causality).

Among other model parameters (either explicit or implicit) reported in Table 11, some are constant over regions, $\boldsymbol{\rho}$, $\boldsymbol{\tau}$, α_1, α_2, $(\lambda_E - \lambda_A)$, $\ln\beta$, ϕ, ρ, σ, d, $(\zeta_E - g_E)$, while others are region-specific, c_i, $\ln\tilde{\alpha}_i$, α_{Ai}, a_i, $\tilde{a}_i$, δ_i. These latter parameters are thus reported as average values across regional observations. For all these structural parameters standard errors are computed through the boostrap procedure discussed in section 5.2.[70]

$\boldsymbol{\tau}$ expresses the structural effect of CAP dummy variables on endogenous variables, while $\boldsymbol{\rho}$ gathers the lagged effects of endogenous variables and thus, together with the structural simultaneous effects ($\mathbf{A}_0$), establishes the long-term relations. With respect to the CAP, structural parameters eventually indicate that agricultural share is negatively affected by both the introduction of the CAP and its 1992 reform, though the former is a very slight impact, and the latter is not statistically significant. At the same time, relative price is negatively affected by the introduction of the CAP, while positively by the MacSharry reform and this latter effect is indeed the largest in magnitude among these parameters.[71]

---

69 Eventually, the only significant response to relative price exogenous shocks concerns relative price itself, as it is positive, significant and persistent over time thus suggesting and interesting, though not easily interpretable, self-reinforcing price increase mechanism.

70 The respective bias of the bootstrap estimates with respect to the real parameter estimate (Efron, 1982) largely varies between a minimum of 3% and a maximum of more than 30%; these biases are available upon request.

71 It must be noticed, however, that within the adopted SVAR structure the effect of CAP dummies on S_{At} can be also indirect via $\ln \overline{P}_{At}$. That is why the structural effect of exogenous variables on endogenous ones can be more correctly evaluated through long-run multipliers of *model 3*.

About half parameters in $\boldsymbol{\rho}$ are actually not statistically significant. As all parameters expressing the contemporaneous effects are significant, we can conclude that in most cases the lagged effects do not affect the relation among endogenous variables, which is eventually limited to the contemporaneous effects and no significant difference occursbetween the short-run and the longer-term relations. Among the few significant parameters associated to the lagged terms, four refer to diagonal elements, that is, concern own lagged effects. These lagged own effects occur in all endogenous variables, relative price being the only case where both lags significantly affect the current value, but in opposite directions such that these lagged effects apparently tend to counterbalance.[72]

In the case of capital intensity and agricultural share, the own-lagged parameters actually indicate that some self-reinforcing mechanism occurs and this is not fully taken into account by the theoretical model (15d)-(12i). In the two cases, as confirmed by the IRFs in Figure 8, this mechanism eventually operates in opposite directions according to the respective prevalent long-term patterns.

In Table 11, most remaining parameters are not statistically different from 0.[73] In some cases, however, this has itself an interesting economic interpretation. As mentioned in section 3.2.3, we can give α_2 a sort of "public agricultural R&D intensity" interpretation; it is expected to be positive but very close to 0.[74] Here a small positive value (.02) is obtained, though the estimated parameter is not significantly different from 0, thus supporting this interpretation.

Particular attention can be also paid to those parameters eventually generating the technological effect explaining agricultural decline, as expressed in (15d). This effect depends on the combination of $\alpha^s_{AE}\alpha_1$ and $\alpha^s_{AE}\left(\lambda_E - \lambda_A\right)$. As shown in Table 11, as α^s_{AE} is significantly negative, both $\left(\lambda_E - \lambda_A\right)$ and α_1 assume a positive value, but not statistically different from 0. The interpretation is somehow consistent with the expectations: a technological gap between E and A might occur, if the agricultural sector did not benefit also from the positive contribution of public agricultural R&D. Consequently, we can conclude

72 The IRFs in Figure 8, however, indicates that even for $\ln \overline{P}_{At}$ a self-reinforcing mechanism actually operates due to the prevalence of the positive lagged effect, that is of $\ln \overline{P}_{At-1}$.

73 This also holds for the two structural parameters of matrices $\mathbf{A}_1$ and $\mathbf{A}_2$, $\rho_{Pk(-1)}$ and $\rho_{Pk(-2)}$, that should be fixed at 0 according to (24). The model estimate confirms they, in fact, are not statistically different from 0. Also the parameter associated to $\left(c_2 S^2_{Ait} + c_3 S^3_{Ait}\right)$ in equation $\ln k_t$ has been not imposes equal to b^s; its estimate, however, is .548 with a bootstrapped standard error of .201; thus, it is not statistically different from the model direct estimate of b^s.

74 Currently in Italy, this value it is definitely lower than .1% (Esposti and Pierani, 2006).

that these parameters and their combination do not seem to justify a significant and relevant technological gap effect on the decline of agricultural share.

Other interesting evidence emerging from Table 11 concerns the demand side of the model. In particular, as $\frac{Y_{A0}}{\left(Y_{A0}-\gamma_A\right)}=\beta$, $\ln\beta$ indicates how relevant the minimum requirement level of agricultural products is with respect to total agricultural production (consumption). Evidently, we expect β be quite close to 1, therefore $\ln\beta$ close to 0. This is indeed confirmed by the estimated parameter which is, in fact, not statistically different from 0, though very slightly negative.

With respect to the inverse demand function, the estimated ρ is significantly negative but not statistically different from -1. As goods A and E must be substitutes, thus $-1<\rho\neq 0$, this result would suggest a value ranging between -.5 and -1, thus a pretty high elasticity of substitution, i.e. σ. The estimated region-specific distribution parameter shows an average value ranging between .14 and .16, which is fully consistent with the underlying theoretical requirement ($0\leq\delta\leq 1$).

A final comment can be spent on other region-specific terms. To identify these parameters, however, we need some additional assumptions. With particular reference to equation $\ln k_t$ and its region-specific constant term $\widetilde{a}_i$, we fix two parameters entering, d and c, thus allowing for the identification of the other components. d indicates the rate of capital depreciation, so it can be considered as a technical coefficient rather than a parameter to be estimated. Here, in fact, we fix d at .15 which is the average value of depreciation implied by the methodology used for the construction of the capital time series (see Appendix 3). c is an approximation term which combines with an error term (see section 4.1.3); thus, fixing it simply implies a re-definition of the error terms. For simplicity, we fix c at 1, thus $\ln c=0$. A final assumption concerns $\ln\widetilde{\alpha}_i$ as it may be of interest to fix its value and evaluate how other parameters are affected. By recalling that $\sum_{i=1}^{N}\widetilde{\alpha}_i=1+\left(N-1\right)\psi$ and, therefore, $\ln\frac{\sum_{i=1}^{N}\widetilde{\alpha}_i}{N}\cong\ln\psi$, we can directly interpret the average value of $\ln\widetilde{\alpha}_i$ as $\ln\psi$, which, in turn, expresses the non-rival part of national public agricultural R&D. As $0\leq\psi\leq 1$, $\ln\widetilde{\alpha}_i=0$ would imply a totally non-rival R&D, while $\ln\widetilde{\alpha}_i=-.5$ would imply about 2/3 of non-rival R&D; on the contrary, $\ln\widetilde{\alpha}_i=.5$, as well as any positive value, is an inconsistent with respect to the theoretical model, as it would imply that the R&D stock available to any region is greater than the total national R&D stock.

Under these assumptions and the three alternative values of $\ln\widetilde{\alpha}_i$, we can thus compute the other parameters and evaluate their respective consistency. $(\zeta_E - g_E)$ is significant and positive under all alternatives and values are all close to .15. This indicates that, on average, most regional production of sector E (85%) is not consumed within the same region as it is prevalently made of intermediate goods (high g_E) and/or export goods (low ζ_E).

α_{Ai} assumes different signs over the three alternatives. Though the parameter is not significant in all cases, it must be reminded that, from a theoretical point of view, a negative value of α_{Ai} is hardly acceptable, as clear in equation (15a). If we admit that this parameter can only take positive value, then we must exclude the two options $\ln\widetilde{\alpha}_i = 0$ and $\ln\widetilde{\alpha}_i = .5$, which are indeed the less plausible among the three cases. More specifically, it must be $\ln\widetilde{\alpha}_i \leq -.4$, thus suggesting that the non-rival part of national agricultural R&D can not be higher than about 2/3.

[Tables 10-11 here]

[Figure 8 here]

7.3. Decomposition of Agricultural Decline

We may use these SVAR estimates to decompose the variation of agricultural sectoral share, S_{At}, into its three fundamental causes: relative price change (Engel'law effect); capital intensification (Rybczynski effect) and technical change (technological gap effect, in turn expressed by both the deterministic trend and sectoral public R&D, R_{At}).

These three effects are clearly expressed in equation (15d), but it is still not easy to achieve this decomposition by simply looking at the estimated parameters due to simultaneity and dynamics within the model. This problem emerges clearly, for instance, looking at the impact of R_{At}, as its effect on S_{At} is not only expressed by parameter $\alpha_{AE}\alpha_1$. Since R_{At} affects $\ln k_t$ as well, it also influences S_{At} indirectly through $\ln k_t$, and this indirect effect depends on parameters α_2 and δ_{AK}.

In previous studies (Gemmell et al., 2000; Punyasavatsut and Coxhead 2002, p. 20; Martin and Warr, 1994, pp. 231-232; Sun et al., 2007, pp. 187-189) this decomposition is achieved through an ECM representation and according to the estimated long-run relations among endogenous variables. A general evidence emerging from these studies is that factor endowment variation (i.e., the Rybczynski effect in our case) is by large the most important cause of agricultural decline.

Here, we decompose the agricultural share variation firstly using the IRFs reported in Figure 8. These are obtained from orthogonalized errors, thus can be directly interpreted as the overall effect of one endogenous variable on the others. Looking at S_{At} , an impulse response is simply $\frac{\partial S_{At+s}}{\partial \varepsilon_t^j}$, where ε_t^j is the structural (orthogonalized) error of the j-th equation, namely the j-th endogenous variable (Pagan and Robertson, 1998). The cumulative response over an infinite horizon ($s \to \infty$) simply is $\theta_j = \sum_{s=0}^{\infty} \frac{\partial S_{At+s}}{\partial \varepsilon_t^j}$.[75]

Secondly, for the n-th exogenous variable, either t or R_{At} in the present case, the long-term impact on S_{At} can be just obtained from the long-run multipliers of the Wickens-Breusch specification (22), $\phi_n = \frac{\partial S_{At}}{\partial x_t^n}$, whose estimates are reported in table 9. Thus, we can aggregate all contributions to net agricultural share change over a given time period (from 0 to t) as follows:

(25) $\Delta S_{At} - \left(\theta_{S_A}\, \Delta S_{At}\right) = \left(1 - \theta_{S_A}\right)\left(S_{At} - S_{A0}\right) = \theta_{P_A}\, \Delta \ln \overline{P}_{At} + \theta_k\, \Delta \ln k_{it} + \phi_R \Delta \ln R_{At} + \phi_t\, t$

Table 12 reports this decomposition for the whole observed period (1951-2002) and three sub-periods (1951-1967; 1968-1984; 1985-2002), as well as for the average region and the two extreme regions.[76]

Considering the whole period, it clearly emerges that the main driver of agricultural decline is the Rybczynski effect, thus confirming previous studies based on very different national contexts. The declining relative price itself contributes to agricultural decline but its relevance is quite limited as it is, in absolute term, the smallest among the three effects. The technological gap effect is, in fact, larger but acts in the opposite direction as it contributes to increase the agricultural sectoral share. These results largely confirm what already presented in previous sections and seem quite robust over time and across regions, as they are largely correspondent for the two extreme regions and for two subperiods.

---

75 It must be reminded that, for I(0) processes, as in the present case, it must be $\lim_{s \to \infty} \frac{\partial S_{At+s}}{\partial \varepsilon_t^j} = 0$ and, consequently, θ_j must be finite. In calculations of Table 11 (and Table A.2) we stop the summation after 15 years.

76 The region-by-region decomposition is reported in Appendix 4. Evidently, the $\phi_R \ln R_{At}$ component is equal across regions.

Only the last subperiod (1985-2002) presents significant differences. Effects are less stable and assume particularly high values. This can be partially explained by the fact that, once agricultural share falls down 5%, its further decline is relatively less affected by structural and regular processes and more influenced by temporary changes (for instance, price adjustments). This is confirmed by the result of the region with the lowest initial share showing this kind of instability in the second subperiod already.

Nonetheless, we cannot exclude more structural explanations especially concerning relative price and technological change patterns, as the capital intensity effect remains stable also in the last period. With respect to the relative price effect, we can partially advocate the role of the progressive reform of the CAP moving from price support to direct income support. Much more difficult, indeed, is the explanation of the changing (in magnitude and direction) technological gap effect, even because, as for relative price, the two extreme regions move in opposite directions. Any possible further interpretation on this aspect, therefore, requires additional and more detailed investigations.

[Table 12 here]

7.4. Regional Differences

There are two major sources of regional heterogeneity within results presented thus far. Firstly, due to the panel specification, the model inherently contains regional-specific fixed-effects; in other words, the model itself differs across regions. Secondly, though based on the same model parameters, the three major effects driving agricultural decline may differ across regions simply because their drivers, namely $\ln \overline{P}_{At}$ and $\ln k_t$, differ.[77] Consequently, the decomposition of agricultural decline is itself region-specific.

Appendix 4 (Table A.1) details the 20 estimated regions-specific constant terms equation-by-equation. In the case of the relative price equation, no regional dummy is statistical significant, thus we can conclude that there is no relevant regional heterogeneity in this respect. This evidence, however, should not surprise as it was already clear by inspection of Figure 2. Same argument applies to the other two endogenous variables. Figure 3 and 4 already showed remarkable differences across regions and this is confirmed by the region-specific terms of respective equations. Most of them are statistically significant, especially in the S_{At} equation. Even in these equations, however, regional shifters maintain the same sign

77 In absolute terms, the technological gap effect is equal across regions. As percentage on total agricultural decline, however, this effect may obviously differ. As may be noticed in Tables A.1 and A.2, however, the relative price effect itself, in absolute values, is quite homogenous across regions, the main difference being thus given by the capital intensity effect.

in all statistically significant cases. This substantial homogeneity of the individual effects, in sign and magnitude, is confirmed also when we compare Northern, Central and Southern regions: in fact, no appreciable systematic difference can be noticed among these regional sub-groups.

This general impression is also validated by looking at the regional decomposition of agricultural decline (Table A.2). No divergent courses actually emerge, as structural change procceds in all regions along an analogous pattern. Nonetheless, the initial conditions significantly differ and this can have an impact on the way structural change occurs. Map A.1(A) and Table A.2 show how agricultural decline occurred with different intensity simply due to different initial development levels.

In those less developed regions where agriculture still had a major role in 1951, namely Southern and Eastern parts of Italy, the decline has been more intense as well as the role of the driving forces. In relative terms, however, the key factor of decline, that is capital intensity effect, was not necessarily prevalent in these less developed regions. We may indeed notice that most Southern regions show a relatively lower level of capital effect, while this is larger especially in Western regions. More generally, looking at figures in Table A.2, we have the impression that, though some significant regional differences may emerge in unravelling agricultural decline into its components, the conventional North vs. South interpretation it is not particularly helpful in this respect, whereas West vs. East seems indeed more indicative.

8. Concluding remarks

This paper deals with a classic topic of development economics, that is why and how agriculture declines as economic development proceeds. The main purpose of the study, therefore, is to answer this question with respect to the intense growth experienced by Italian regions after the Second World War.

In achieving this main objective, two major aspects are dealt with. Firstly, agricultural decline is modelled within a general equilibrium framework in such a way that all major driving forces involed reciprocally and dynamically interact. Secondly, the empirical approach proposed copes with all the relevant econometric issues implicated by the dynamic and panel data representation of such simultaneous equation model. Model estimation is thus achieved through a sequence of steps and alternative specifications eventually leading to the identification of structural relations among the variables and, consequently, the decomposition and attribution of agricultural decline to its major drivers.

The most relevant result actually concerns the relation among agricultural declining share and capital intensification within the economy. While the effect of the former on the latter, though positive, does not reveal a key role of the sector in the growth process, capital intensification is by large and in all regions the major driver of agricultural decline, while the relative price effect is much smaller and the technological gap effect actually plays in favour of an higher agricultural share.

Not only this result confirms most previous analogous studies, but seems largely shared by all the Italian regions, without any significant difference between the Northern and the Southern part of the country. With respect to the conventional interpretation of the Italian unequal regional development, more interesting is indeed the difference emerging between Western and Eastern regions.

A less homogenous picture emerges with reference to the last subperiod, from mid-eighties to the new century. This can be interpreted by observing that, when most structural change involving agriculture is indeed completed, its further decline and, more generally, its interaction with overall economic development may be actually questioned or assume different characters and drivers. On this aspect, however, further research effort should be spent in the future especially focusing on the role of agriculture within highly-developed and post-industrial economies.

References

Aasness, J., Biørn, E., Skjeper, T. 1993. Engel functions, panel data and latent variables. *Econometrica* 61: 1395-1422.

Abbott, P., McCalla, A. 2002. Agriculture in the Macroeconomy: Theory and Measurement. In: Gardener, B.L., Rausser, G.C. (eds.), In Gardner, B., Rausser, G. (eds.), *Handbook of Agricultural Economics, Volume 2A.* Amsterdam: Elsevier Science, 1659-1686.

Ahammad, H. 2002. A CGE Approach to Measuring the Sectoral Contribution to an Economy: An Application to Western Australian Agriculture. *Review of Urban & Regional Development Studies* 14(3), 305-324.

Amisano, G., Giannini, C. 1997. *Topics in Structural VAR Econometrics*. 2nd Edition. Heildelberg: Springer.

Anderson, K. 1987. On why agriculture declines with economic growth. *Agricultural Economics* 1, 195–207.

Arellano, M. 2003. *Panel-data Econometrics*. Oxford: Oxford University Press.

Arellano, M., Bond, S. 1991. Some Tests of Specification for Panel-data: Monte Carlo Evidence and an Application to Employment Equations. *Review of Economic Studies* 58, 277-297.

Ball, V.E., Bureau, J.C., Butault, J.P., Witzke, H.P. 1993. The Stock of Capital in European Community Agriculture. *European Review of Agricultural Economics* 20, 437-50.

Baltagi, B.H. 2005. *Econometric Analysis of Panel Data*. New York: John Wiley & Sons.

Barro, R. 1991. Economic Growth in a Cross Section of Countries. *Quarterly Journa/ of Economics* 106(2), 407-433.

Beckman, S.R., Smith, W.J. 1993. Positively Sloping Marginal Revenue, CES Utility and Subsistence Requirements. *Southern Economic Journal*, 60(2), 297-303.

Blundell, R., Bond, S. 1998. Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics* 87, 115-143.

Chambers, R.G. 1988. *Applied production analysis. A dual approach*. Cambridge: Cambridge University Press.

Chernick, M. R. 1999. *Bootstrap Methods: A Practitioner's Guide.* New York: Wiley.

Cornwell, C., Schmidt, P., Wyhowki, D. 1992. Simultaneous equations and panel data. *Journal of Econometrics* 51, 151-181.

Davison, A.C., Hinkley, D.V. 1997. *Bootstrap Methods and their Application*. Cambridge: Cambridge University Press.

Deaton, A., Mullenbauer, J. 1980. *Economics and consumer behaviour*. Cambridge: Cambridge University Press.

Dhrymes, P.J. 2006. *Econometric Structure of Panel Data Models*. Columbia University, Unpublished Paper.

Diao, X., Hazell, P., Resnick, D., Thurlow, J. 2007. *The Role of Agriculture in Development. Implications for Sub-Saharan Africa*. Research Report n. 153, International Food Policy Research Institute (IFPRI), Washington D.C..

Diewert, E.W. 2005. Issues in the Measurement of Capital Services, Depreciation, Asset Price Changes and Interest Rates. In: Corrado, C., Haltiwanger, J., Sichel D. (eds.), *Measuring Capital in the New Economy*, Chicago: University of Chicago Press, 479-542.

Echevarria, C. 1997. Changes in Sectoral Composition Associated with Economic Growth. *International Economic Review* 38, 431-452.

Efron, B. 1982. *The Jackknife, the Bootstrap and Other Resampling Plans*. Philadelphia: Society for Industrial and Applied Mathematics.

Enders, W. 1995. *Applied Econometric Time Series*. New York: John Wiley & Sons.

Esposti, R. 2007. Regional growth and policies in the European Union: Does the Common Agricultural Policy have a counter-treatment effect? *American Journal of Agricultural Economics* 89 (1), 116-134.

Esposti, R., Pierani, P. 2000. Modelling Technical change in Italian Agriculture: a Latent Variable Approach. *Agricultural Economics* 22 (3), 261-270.

Esposti, R., Pierani, P. 2003. Building the Knowledge Stock: Lags, Depreciation, and Uncertainty in R&D Investment and Productivity Growth. *Journal of Productivity Analysis* 19 (1), 33-58.

Esposti, R., Pierani, P. 2006. Price, Private Demand and Optimal Provision of Public R&D Investment. An application to Italian agriculture, 1960-1995. *Empirical Economics* 31 (3), 699-715.

Foellmi, R., Zweimuller, J. 2006. Structural Change, Engel's Consumption Cycles and Kaldor's Facts of Economic Growth. Institute for Empirical Research in Economics, University of Zurich, Unpublished Paper.

Fuà, G. 1992. Il declino dell'agricoltura. In: G. Fuà (ed.), *Lo Sviluppo economico in Italia. Lavoro e reddito*. Milano: Franco Angeli, 140-155.

Fulginiti, L., Onofri, A. 2006. *Public Inputs and Dynamic Producer Behavior: Endogenous Growth in U.S. Agriculture*. Univerity of Nebraska, Agricultural Economics Department, Lincoln.

Gardner, B. 2003. *Causes of Rural Economic Development*. Elmhirst Lecture at the 25th Conference of the IAAE, August 17, Durban, South Africa.

Gemmell, N., Lloyd, T.A., Mathew, M. 2000. Agricultural Growth and Inter-Sectoral Linkages in a Developing Economy. *Journal of Agricultural Economics* 51(3), 353-370.

Gollin, D., Parente, S., Rogerson, R. 2002. The Role of Agriculture in Development. *American Economic Review* 92(2), 160-164.

Greenwood, J., Seshadri, A. 2001. Technological Progress and Economic Transformation. In: Aghion, P., Durlauf, S.N. (eds.), *Handbook of Growth Economics. Volume IB*. Amsterdam: Elsevier, 1227-1273.

Greene, W.H. 1997. *Econometric Analysis. 3rd Edition*. New York: Prentice-Hall.

Hansen, G.H., Prescott, E.C. 2002. Malthus to Solow. *American Economic Review* 92(4), 1205-1217.

Hendry, D.F., Pagan, A.R., Sargan, J.D. 1984. Dynamic specification. In: Griliches, Z., Intriligator, M. (eds.), *Handbook of Econometrics. Vol. II*. Amsterdam: North Holland.

Hendry, D.F. 1995. *Dynamic Econometrics*. New York: Oxford University Press.

Holtz-Eakin, D., Newey, W., Rosen, H. 1988. Estimating vector autoregressions with panel data. *Econometrica* 56, 1371-1395.

Hsiao, C. 2003. *Analysis of Panel Data. 2nd Edition*. Cambridge: Cambridge University Press.

Hulten, C.R. 1990. The Measurement of Capital. In: Berndt, E.R., Triplett, J.E. (eds.), *Fifty Years of Economic Measurement (The Jubilee of the Conference on Research in Income and Wealth*), NBER: Studies in Income and Wealth, Chicago: The University of Chicago Press, 119-52.

Im, K.S., Pesaran, M.H., Shin, Y. 2003. Testing for unit roots in heterogeneous panels. *Journal of Econometrics* 115, 53-74.

Irz, X., Roe, T. 2005. Seeds of growth? Agricultural productivity and the transitional dynamics of the Ramsey model. *European Review of Agricultural Economics* 32(2), 143-165.

Jouini, J. 2006. *Bootstrap Tests in Bivariate VAR Process with Single Structural Change: Power versus Corrected Size and Empirical Illustration*. Working Paper 2006-41, GREQAM, Université de la Méditerranée, Marseille.

Kohli, U. 1978. Gross national product function and the derived demand for imports and exports. *Canadian Journal of Economics* 11(2), 167–182.

Kohli, U. 1991. *Technology, Duality, and Foreign Trade: The GNP Function Approach to Modeling Imports and Exports*. University of Michigan Press, Ann Arbor, MI.

Kim, H.Y. 1997. Inverse Demand Systems and Welfare Measurement in Quantity Space. *Southern Economic Journal*, 63(3), 663-79.

Kongsamut, P. Rebelo, S., Xie, D. 2001. Beyond Balanced Growth. *The Review of Economic Studies*, 68, 869-882.

Laitner, J. 2000. Structural Change and Economic Growth. *The Review of Economic Studies*, 67, 545-561.

Lastrapes, W.D., McMillin, W.D. 2004. Cross-Country Variation in the Liquidity Effect: The Role of Financial Markets. *The Economic Journal* 114, 890-915.

Lewis, W.A. 1954. Economic development with unlimited supplies of labor. *Manchester School of Economic and Social Studies* 22, 139–191.

Lucchetti, R. 2006. Identification of Covariance Structures. *Econometric Theory* 22, 235-257.

Lundberg, M., Squire, L. 2003, The Simultaneous Evolution of Growth and Inequality. *The Economic Journal* 113, 326-344.

Lupi, C., Mantegazza, S. 1994. *Ricostruzione delle serie degli investimenti per branca utilizzatrice, per branca proprietaria e calcolo dello stock di capitale*. Quaderni di Ricerca n. 14. Roma: ISTAT.

Lütkepohl, H. 1991. *Introduction to Multiple Time Series Analysis*. 1st Edition. New York: Springer.

Lütkepohl, H. 2000. *Bootstrapping Impulse Response in VAR Analyses*. Working Paper 2000-22, Sonderforschungsbereich 373, Humboldt University, Berlin.

Maietta, O.W. 2004. Crescita della produttività e capitale umano nell'agricoltura italiana. *La Questione Agraria*, 2004/n.4, 105-140.

Martin, W., Warr, P.G. 1993. Explaining the relative decline of agriculture: a supply-side analysis for Indonesia. *World Bank Economic Review* 7, 381–401.

Martin, W., Warr, P.G. 1994. Determinants of agriculture's relative decline: Thailand. *Agricultural Economics* 11, 219–235.

Mundlak, Y. 2000. *Agriculture and Economic Growth.* Cambridge: Harvard University Press.

Ngai, L.R., Pissarides, C.A. 2007. Structural Change in a Multisector Model of Growth. *American Economic Review* 97, 429-443.

Omtzigt, P., Fachin, S. 2002. *Bootstrapping and Bartlett corrections in the cointegrated VAR model*. Working Paper 2002-24, Facoltà di Economia - Università dell'Insubria, Varese.

Paci, R., Saba, A. 1997. *The empirics of regional economic growth in Italy. 1951-93*. Contributi di Ricerca CRENoS n. 97/1, Università degli Studi di Cagliari.

Paci, R., Pusceddu, N. 1999. *Stima dello stock di capitale nelle regioni italiane. 1970-1994*. Contributi di Ricerca CRENoS n. 99/5, Università degli Studi di Cagliari.

Pagan, A.R., Robertson, J.C. 1998. Structural models of the liquidity effect. *The Review of Economics and Statistics* 80(2), 202-217

Primont, D.F., Primont, D. 1995. Further Evidence of Positively Sloping Marginal Revenue. *Southern Economic Journal*, 62(2): 481-85.

Punyasavatsut, C., Coxhead, P. 2002. *On the Decline of Agriculture in Developing Countries: A Reintepretation of the Evidence*. Staff Paper Series n. 457, Department of Agricultural & Applied Economics, University of Wisconsin at Madison.

Rizzi, P.L., Pierani P. 2006. *AGREFIT. Ricavi, costi e produttività dei fattori nell'agricoltura delle regioni italiane (1951-2002)*. Milano: Franco Angeli.

Robinson, S. 2001. Multisectoral Models. In: Chenery, H., Srinivasan, T.N. (eds.), Handbook *of Development Economics. Volume II.* Amsterdam: Elsevier Science, 885-947.

Romer, P.M. 1986. Increasing returns and long run growth. *Journal of Political Economy* 94, 1002–37.

Romer, P.M. 1990. Endogenous Technological Change. *Journal of Political Economy* 98, S71-S102.

Roodman, D. 2006. *How to Do xtabond2: An Introduction to "Difference" and "System" GMM in Stata*. Washington D.C.: Center for Global Development, Working Paper n. 103.

Self, S., Grabowski, R. 2007. Economic development and the role of agricultural technology. *Agricultural Economics* 36, 395-404.

Solow, R.M. 1956. A contribution to the theory of economic growth. *Quaterly Journal of Economics* 70, 65-94.

Stock, J.H., Watson, M.W. 2001. Vector Autoregressions. *Journal of Economic Perspectives* 15 (4), 101-115.

Strickland, A., Weiss, L. 1976. Advertising, Concentration, and Price Cost Margins. *Journal of Political Economy* 84, 1109-1121.

Sun, L., Fulginiti, L.E., Peterson, E.W.F. 2007. Accounting for agricultural decline with economic growth in Taiwan. *Agricultural Economics* 36, 181–190.

Swensen, A.R. 2006. Bootstrap Algorithms for Testing and Determining the Cointegration Rank in VAR Models. *Econometrica* 74(6), 1699-1714.

Syrquin, M. 1988. Patterns of Structural Change. In: Chenery, H., Srinivasan, T.N. (eds.), *Handbook of Development Economics. Volume I.* Amsterdam: Elsevier Science, 203-273.

Takayama, A. 1965. On a two-sector model of economic growth with technological change. *Review of Economic Studies* 32, 251-262.

Tiffin, R., Irz, X. 2006. Is agriculture the engine of growth? *Agricultural Economics* 35, 79-89.

Timmer, C.P. 1988. The Agricultural Transformation. In: Chenery, H., Srinivasan, T.N. (eds.), *Handbook of Development Economics, vol. I.* Amsterdam: Elsevier Science, 275-331.

Timmer, C.P. 2002. Agriculture and Economic Development. In: Gardner, B., Rausser, G. (eds.), *Handbook of Agricultural Economics, Volume 2A.* Amsterdam: Elsevier Science, 1487-1546.

Uzawa, H. 1961. On a two-sector model of economic growth, I. *Review of Economic Studies* 29, 40-47.

Uzawa, H. 1963. On a two-sector model of economic growth, II. *Review of Economic Studies* 30,105-118.

Uzawa, H. 1964. Optimal growth in a two-sector model of capital accumulation. *Review of Economic Studies* 31, 1-24.

Vallanti, G. 2007. *Capital Mobility and Unemployment Dynamics: Evidence from a Panel of OECD Countries*. Working Document n. 42, LLEE-Luiss "Guido Carli", Rome.

Veerbek, M. 2004. *A Guide to Modern Econometrics. 2nd Edition*. New York: John Wiley & Sons.

Weisdorf, J. 2006. From Domestic Manufacture to Industrial Revolution: Long-Run Growth and agricultural Development. *Oxford Economic Papers* 58(2), 264-287.

Wickens, M.R., Breusch, T.S. 1988. Dynamic specification, the long-run and the estimation of transformed regression models. *The Economic Journal* 98, 189-205.

Windmeijer, F. 2005. A Finite Sample Correction for the Variance of Linear Efficient Two-Step GMM Estimators. *Journal of Econometrics* 126(1), 25-51.

Wong, K.K.G., McLaren, K.R. 2005. Specification and Estimation of Regular Inverse Demand Systems: A Distance Function Approach. *American Journal of Agricultural Economics* 87(4): 823–834.

Zellner, A., Palm, F.C. 1974. Time series analysis and simultaneous equation econometric models. *Journal of Econometrics* 2, 17-54.

Zellner, A., Palm, F.C. 2004. *The Structural Econometric Time Series Analysis Approach*. Cambridge: Cambridge University Press.

Table 1 – Descriptive statistics of model variables (20 regions; period 1951-2002)

	Endogenous			**Exogenous**	
	$\overline{P}_{At}$	k_t	S_{At}	y_{At}	R_{At}
Initial (1951) values					
Mean	2.23	9.12	0.196	1.09	91.65
Max.	2.77	17.47	0.365	1.95	91.65
Min.	1.84	2.82	0.067	0.60	91.65
CV	0.11	0.42	0.48	0.27	0.00
Final (2002) values					
Mean	0.89	85.18	0.027	1.91	2349.47
Max.	0.95	114.70	0.055	3.02	2349.47
Min.	0.79	66.05	0.011	0.92	2349.47
CV	0.04	0.15	0.42	0.33	0.00
Avg. annual growth rates (02-51)					
Mean	-0.018	0.046	-0.037	0.011	0.066
Max.	-0.014	0.070	-0.027	0.024	0.066
Min.	-0.023	0.031	-0.045	0.003	0.066
CV	-0.137	0.184	-0.144	0.571	0.000

Table 2 – Panel unit root test (IPS)

	1 lag		2 lags	
Variable	With trend	Without trend	With trend	Without trend
$\ln \overline{P}_{At}$	3.48 (.001)	8.98 (.000)	3.33 (.001)	8.55 (.000)
$\ln k_t$	-5.94 (.000)	-11.57 (.000)	-5.32 (.000)	-9.40 (.000)
S_{At}	1.95 (.050)	-5.82 (.000)	2.92 (.004)	-5.66 (.000)
$\ln y_{At}$	-4.48 (.000)	-4.32 (.000)	-4.31 (.000)	-4.50 (.000)
$^{*}\ln R_{At}$	-2.79 (.199)	-1.32 (.618)	-0.585 (.970)	-1.67 (.441)

Note: asymptotic p-values in parenthesis below the estimated values
Software: TSP4.5
** Conventional ADF unit root test (52 observations)*

Table 3 – Unrestricted VAR model OLS estimates (equation-by-equation): model 1

Equation	$\ln \overline{P}_{At}$			$\ln k_t$			S_{At}		
	OLS (Static)	*OLS (Pooled)*	*LSDV (Within)*	*OLS (Static)*	*OLS (Pooled)*	*LSDV (Within)*	*OLS (Static)*	*OLS (Pooled)*	*LSDV (Within)*
Variable									
$\ln \overline{P}_{At-1}$	-	.986*	.982*	-	.027	.026	-	.071*	.070*
		(.032)	(.032)		(.019)	(.019)		(.009)	(.008)
$\ln k_{t-1}$	-	-.053	-.040	-	1.318*	1.275*	-	.018	.001
		(.051)	(.053)		(.030)	(.030)		(.014)	(.014)
S_{At-1}	-	-.383*	-.406*	-	.036	.013	-	.335*	.321*
		(.095)	(.098)		(.055)	(.056)		(.025)	(.026)
$\ln \overline{P}_{At-2}$	-	-.071*	-.077*	-	-.030	-.037*	-	-.068*	-.066*
		(.033)	(.033)		(.019)	(.019)		(.009)	(.009)
$\ln k_{t-2}$	-	.069	.073	-	-.340*	-.312*	-	-.024*	-.021
		(.050)	(.050)		(.029)	(.029)		(.012)	(.013)
S_{At-2}	-	.411*	.403*	-	-.001	-.024	-	.561*	.535*
		(.093)	(.096)		(.054)	(.055)		(.025)	(.025)
$\ln y_{At}$	-.036	-.002	-.029*	.622*	.001	.005	-.030*	.005*	.020*
	(.020)	(.005)	(.010)	(.042)	(.003)	(.006)	(.006)	(.001)	(.003)
$\ln R_{At}$	.093*	.002	-.008	.564*	.013*	.020*	-.016*	.002	.007*
	(.019)	(.009)	(.010)	(.040)	(.005)	(.006)	(.006)	(.003)	(.003)
t	-.025*	-.003*	-.003*	-.011*	-.002*	-.002*	-.002*	-.000	.000
	(.001)	(.001)	(.001)	(.003)	(.001)	(.001)	(.001)	(.000)	(.000)
CAP1	.084*	.013*	.013*	.346*	-.004	.001	-.010*	-.006*	-.003*
	(.013)	(.006)	(.006)	(.027)	(.003)	(.004)	(.004)	(.002)	(.001)
CAP2	-.099*	.022*	.025*	-.136*	.010*	.008*	.018*	-.001	.002
	(.013)	(.006)	(.007)	(.027)	(.004)	(.004)	(.004)	(.002)	(.002)
Adj. R^2	.911	.981	.981	.929	.999	.999	.840	.974	.975
AIC	-1022	-1759	-1770	-254	-2313	-2316	-2208	-3084	-3103

Note: Estimated standard errors in parenthesis below the estimated values
Software: TSP4.5
** Statistically significant at the 5% level*

Table 4 – Unrestricted VAR model GMM estimates (equation-by-equation): model 1

Equation	$\ln \overline{P}_{At}$		$\ln k_t$		S_{At}	
	GMM one-step	*GMM 2-teps*	*GMM one-step*	*GMM 2-teps*	*GMM one-step*	*GMM 2-teps*
Variable						
$\ln \overline{P}_{At-1}$	*.981** *(.0324)*	*1.054** *(.133)*	*.024* *(.019)*	*.013* *(.041)*	*.064** *(.008)*	*.045** *(.015)*
$\ln k_{t-1}$	*-.028* *(.056)*	*.087* *(.219)*	*1.269** *(.031)*	*1.188** *(.119)*	*.002* *(.014)*	*.018* *(.032)*
S_{At-1}	*-.533** *(.103)*	*-.776* *(.670)*	*.018* *(.058)*	*.086* *(.357)*	*.309** *(.025)*	*.342** *(.077)*
$\ln \overline{P}_{At-2}$	*-.072** *(.034)*	*-.139* *(.083)*	*-.037** *(.018)*	*.000* *(.041)*	*-.059** *(.009)*	*-.054** *(.009)*
$\ln k_{t-2}$	*.058* *(.052)*	*-.064* *(.178)*	*-.306** *(.029)*	*-.261** *(.090)*	*-024** *(.012)*	*-.039* *(.027)*
S_{At-2}	*.366** *(.101)*	*.386* *(.312)*	*-.019* *(.057)*	*-.260* *(.476)*	*.515** *(.025)*	*.504** *(.059)*
$\ln y_{At}$	*-.029** *(.011)*	*-.016* *(.080)*	*.002* *(.007)*	*.011* *(.044)*	*.024** *(.002)*	*.044** *(.013)*
$\ln R_{At}$	*-.010* *(.010)*	*-013* *(.030)*	*.211** *(.006)*	*.024* *(.019)*	*.008** *(.003)*	*.004* *(.008)*
t	*-.004** *(.001)*	*-.002* *(.004)*	*-.002** *(.000)*	*-.001* *(.002)*	*-.000* *(.000)*	*-.000* *(.000)*
CAP1	*.011* *(.007)*	*-.002* *(.019)*	*.002* *(.004)*	*.000* *(.003)*	*-.004** *(.002)*	*-.007** *(.002)*
CAP2	*.027** *(.007)*	*-1.185* *(2.051)*	*.008** *(.004)*	*.004* *(.004)*	*-.003* *(.002)*	*-.003** *(.001)*
LM test 1st order autocorrelation	*-21.20** *(.000)*	*-2.380** *(.017)*	*-17.01** *(.000)*	*-3.100** *(.002)*	*-18.190** *(.000)*	*-2.750** *(.006)*
LM test 2nd order autocorrelation	*.040* *(.967)*	*3.040** *(.002)*	*.380* *(.703)*	*.060* *(.951)*	*-7.390** *(.000)*	*-1.640* *(.100)*
Sargan test of overidentifying restrictions	*890.35* *(1.000)*	*17.79* *(1.000)*	*902.64* *(1.000)*	*18.28* *(1.000)*	*909.36* *(1.000)*	*15.59* *(1.000)*

Note: Estimated standard errors in parenthesis below the estimated values; p-values in parenthesis under the test values
Software: STATA9
** Statistically significant at the 5% level*

Table 5 – Restricted VAR model SUR estimates: model 2

Equation / *Variable*	$\ln \overline{P}_{At}$	$\ln k_t$	S_{At}
$\ln \overline{P}_{At-1}$	*.964** (*.033*)	*-.124* (*.102*)	*.064** (*.008*)
$\ln k_{t-1}$	*.000*	*.891** (*.164*)	*-.001* (*.013*)
S_{At-1}	*-.444** (*.097*)	*.508* (*.318*)	*.318** (*.026*)
$\ln \overline{P}_{At-2}$	*-.048* (*.033*)	*.274** (*.103*)	*-.063** (*.008*)
$\ln k_{t-2}$	*.000*	*-.343** (*.159*)	*-.012* (*.013*)
S_{At-2}	*.309** (*.097*)	*-1.245** (*.305*)	*.532** (*.025*)
$\ln y_{At}$	*.000*	*1.000*	*.000*
$\ln R_{At}$	*.000*	*.067** (*.030*)	*.008** (*.002*)
T	*-.003** (*.001*)	*.002* (*.002*)	*-.001** (*.000*)
CAP1	*.027** (*.005*)	*.000*	*-.002** (*.001*)
CAP2	*.023** (*.006*)	*.000*	*.001* (*.001*)
Adj. R^2	*.980*	*.954*	*.974*
AIC		*-5408***	

Note: Estimated standard errors in parenthesis below the estimated values
Software: TSP4.5
** Statistically significant at the 5% level*
*** Referring to the whole system of equations*

Table 6 – Restricted VAR model joint-GMM estimates: model 2

Equation	$\ln \overline{P}_{At}$		$\ln k_t$		S_{At}	
	J-GMM one-step	*J-GMM two-steps*	*J-GMM one-step*	*J-GMM two-steps*	*J-GMM one-step*	*J-GMM two-steps*
Variable						
$\ln \overline{P}_{At-1}$	.980* (.033)	1.031* (.068)	-.138* (.065)	-.207 (.173)	.066* (.009)	.052* (.015)
$\ln k_{t-1}$	.000	.000	.415* (.037)	.375* (.100)	.006 (.015)	.021 (.016)
S_{At-1}	-.588* (.101)	-.796* (.274)	.617 (.275)	-.319 (.873)	.295* (.027)	.320* (.060)
$\ln \overline{P}_{At-2}$	-.061 (.034)	-.079 (.052)	.204* (.065)	.187 (.200)	-.063* (.009)	-.060* (.007)
$\ln k_{t-2}$	.000	.000	.408 (.035)	.329* (.135)	-.022 (.014)	-.035* (.014)
S_{At-2}	.278* (.098)	.374 (.325)	-.605* (.267)	.108 (.672)	.530* (.026)	.516* (.048)
$\ln y_{At}$	.000	.000	1.000	1.000	.000	.000
$\ln R_{At}$	.000	.000	.057* (.017)	.071 (.094)	.008* (.003)	.011* (.005)
t	-.004* (.001)	-.002 (.002)	.004* (.001)	-.002 (.005)	-.000 (.000)	-.001 (.000)
CAP1	.016* (.005)	.006 (.010)	.000	.000	-.004* (.002)	-.005* (.002)
CAP2	.024* (.006)	-.037 (.070)	.000	.000	.001 (.001)	-.004* (.001)

	one-step	*two-steps*
*LM test 1st order autocorrelation***	-22.07* (.000)	-3.200* (.001)
*LM test 2nd order autocorrelation***	.120 (.902)	.470 (.636)
*Sargan test of overidentifying restrictions***	964.29 (1.000)	18.18 (1.000)

Note: J-GMM=joint GMM estimation
Estimated standard errors in parenthesis below the estimated values
Software: STATA9
** Statistically significant at the 5% level*
*** Referring to the whole system of equations*

Table 7 – Lags and specification selection criteria (SUR estimation)

	Pooled (2-lags)	*1-lag*	*2-lags*	*3-lags*	*Restricted (2-lags)*
Log Likelihood	*7230*	*7150*	*7300*	*7208*	*5491*
AIC	*-7194*	*-7066*	*-7207*	*-7106*	*-5408*

Note: values refer to the on the whole system of equations

Table 8 – Granger-causality: F-tests

	F-block exogeneity	*Regressor* $\ln \overline{P}_{At}$	$\ln k_t$	S_{At}
Dependent variable				
$\ln \overline{P}_{At}$	*9.213 (.01)*	-	*1.290 (.63)*	*10.769 (.01)*
$\ln k_t$	*1.181 (.32)*	*2.343 (.39)*	-	*.455 (.88)*
S_{At}	*35.093 (.01)*	*42.002 (.01)*	*3.958 (.22)*	-

Note: p-values in parenthesis
Tests refer to the unrestricted VAR LSDV estimate (2 lags)

Table 9 – 3SLS estimate of (restricted) Wickens-Breusch specification: model 3

Equation / *Variable*	$\ln \overline{P}_{At}$	$\ln k_t$	S_{At}
$\Delta_1 \ln \overline{P}_{At}$	-4.789 (3.728)	-13.454* (5.828)	-1.725* (.210)
$\Delta_1 \ln k_t$	9.620* (7.917)	1.758 (14.499)	.589 (2.168)
$\Delta_1 \ln S_{At}$	5.699 (10.101)	-15.685 (18.501)	2.262 (2.767)
$\Delta_2 \ln \overline{P}_{At}$	.121 (2.013)	8.313* (3.687)	1.309* (.551)
$\Delta_2 \ln k_t$	-7.999* (3.174)	-10.337 (15.276)	-.487 (1.144)
$\Delta_2 \ln S_{At}$	-14.181* (7.340)	-6.296 (15.276)	-9.821* (2.284)
$\Delta_1 \ln y_{At}$	-.938 (1.256)	.678 (2.300)	-.284 (.344)
$\Delta_1 \ln R_{At}$	2.456* (0.266)	3.397 (2.319)	.747* (.346)
$\Delta_2 \ln y_{At}$	1.256 (.973)	1.090 (1.781)	1.060 (.266)
$\Delta_2 \ln R_{At}$	-.925 (.636)	-2.009* (0.164)	-.316* (.154)
$\ln y_{At}$	.000	1.000	.000
$\ln R_{At}$	.000	1.176* (.075)	.004* (.002)
t	-.059* (.009)	-.099* (.016)	.002 (.002)
CAP1	.154* (.074)	.000	-.060* (.026)
CAP2	.295* (.119)	.000	-.030 (.033)
Adj. R^2	.356	.408	.370

Note: Estimated standard errors in parenthesis below the estimated values
Software: TSP4.5
** Statistically significant at the 5% level*

Table 10 – SVAR estimate (model 4; short-run identifying restrictions): $\mathbf{A}_0$ *(MLE) and restricted VAR (ISUR)*

$\mathbf{A}_0$			
α^s_{AE}		-5.403* (.128)	
∂^s_{AK}		-92.950* (2.077)	
b^s		.335* (0.024)	
$\mathbf{C}_1, \mathbf{C}_2, \mathbf{D}_0$	$\ln \overline{P}_{At}$	$\ln k_t$	S_{At}
$\ln \overline{P}_{At-1}$	1.074* (.037)	.028 (.018)	.070* (.008)
$\ln k_{t-1}$	-.065 (.059)	1.279* (.030)	-.001 (.013)
S_{At-1}	0,073 (.116)	0.013 (.055)	.321* (.025)
$\ln \overline{P}_{At-2}$	-.166* (0.038)	-.038* (.018)	-.063* (.008)
$\ln k_{t-2}$	.080 (.057)	-.313* (.029)	-.017 (.013)
S_{At-2}	1.195* (.122)	-.018 (.053)	.536* (.024)
$\ln y_{At}$	.000	.020* (.003)	.000
$\ln R_{At}$	.000	.020* (.006)	.007* (.001)
$\left(c_2 \mathbf{S}^2_{\mathbf{A}t-2} + c_3 \mathbf{S}^3_{\mathbf{A}t-2}\right)$	.059* (0.004)	.012 (.011)	.000
t	-.002* (.001)	-.002* (.001)	-.001 (.000)
CAP1	.007 (.005)	.000	-.005* (.001)
CAP2	022* (.006)	.000	.008* (.003)
Adj. R^2	.975	.999	.976

Note: Estimated standard errors in parenthesis below the estimated values
Software: STATA9
** Statistically significant at the 5% level ** Referring to the whole system of equations*

Table 11 – SVAR indirect estimates: other explicit and implicit structural parameters (bootstrapped standard errors in parenthesis)

Other explicit parameters:

$\mathbf{A}_1$		$\mathbf{A}_2$		$\mathbf{B}_0^1, \mathbf{B}_0^2$	
ρ_1	.321 (1.490)	ρ_9	39.724* (16.194)	$\alpha_{AE}^s \alpha_1$	-.047* (.021)
ρ_2	35.298* (6.732)	ρ_{10}	-5.539 (4.639)	α_2	.020 (.021)
ρ_3	1.279 (.898)	ρ_{11}	-.313 (1.319)	$\alpha_{AE}^s(\lambda_E - \lambda_A)$	-.061 (.063)
ρ_4	29.837* (9.529)	ρ_{12}	51.035* (16.033)	$-b^s \ln\beta$	.011 (.026)
ρ_5	7.645* (2.863)	ρ_{13}	-5.796* (2.501)	ϕ	.017 (.015)
ρ_6	.000 (.539)	ρ_{14}	-13.131* (5.125)	τ_1	-.005* (.002)
ρ_7	-11.803* (3.328)	ρ_{15}	1.784 (1.281)	τ_2	-.021 (.030)
ρ_8	1.279* (.265)	ρ_{16}	-.313 (.301)	τ_3	-.428* (.162)
$\rho_{Pk(-1)}$	-6.907 (3.945)	$\rho_{Pk(-2)}$	1.690 (1.338)	τ_4	+.733* (.302)

Implicit parameters:

		Region-specific Parameters (average values over 20 regions) according to the fixed value of $\ln\tilde{\alpha}_i$:			
α_1	.005 (.210)				
$\ln\beta$	-.033 (.160)	$\ln\tilde{\alpha}_i$	0 (fixed)	.5 (fixed)	-.5 (fixed)
$(\lambda_E - \lambda_A)$	.011 (.763)	α_{Ai}	-.010 (.008)	-.024 (.017)	.003 (.006)
ρ	-1.251* (.551)	a_i	1.529* (.389)	1.529* (.389)	1.529* (.389)
d	.15 (fixed)	$\tilde{a}_i$	-.005 (.005)	-.015 (.014)	.005 (.007)
c_i	1 (fixed)	δ_i	.015 (.061)	.014 (.060)	.015 (.063)
		$(\zeta_E - g_E)$	.157* (.005)	.147* (.015)	.167* (.007)

Note: Standard errors are computed with the bootstrap method described at section 5.2 and 1000 replications
Software: STATA9
** Statistically significant at the 5% level (Normal distribution assumed)*

Table 12 – Decomposition of agricultural share variation

	Total observed ΔS_{At}	Own effect $\theta_{S_A}\,\Delta S_{At}$	*Components (equation (25))* Relative price effect $\theta_{P_A}\,\Delta \ln \overline{P}_{At}$	Capital intensity effect $\theta_k \Delta \ln k_{it}$	Technological gap effect $\phi_R \Delta \ln R_{At} + \phi_t\, t$
Period: 1951-2002					
Avg. Region	-.1685	.0083	16.6%	116.0%	-32.6%
Highest S_A Region	-.3355	.0165	1.9%	109.1%	-2.0%
Lowest S_A Region	-.0561	.0028	19.5%	117.5%	-37.0%
Period: 1951-1967					
Avg. Region	-.0790	.0039	6.2%	108.0%	-14.2%
Highest S_A Region	-.1030	.0051	5.2%	105.1%	-1.3%
Lowest S_A Region	-.0323	.0016	5.3%	108.8%	-14.1%
Period: 1968-1984					
Avg. Region	-.0643	.0032	14.9%	141.3%	-56.2%
Highest S_A Region	-.1850	.0091	3.8%	115.4%	-19.2%
Lowest S_A Region	-.0145	.0007	198.4%	74.6%	-839.0%
Period: 1985-2002					
Avg. Region	-.0252	.0012	529.5%	195.3%	-624.8%
Highest S_A Region	-.0475	.0023	-348.4%	83.7%	364.8%
Lowest S_A Region	-.0093	.0005	133.6%	10.2%	-133.8%

Figure 2 – Evolution of $\overline{\mathbf{P}}_{\mathbf{A}t}$ over the sample period: comparison between the average and the extreme regions (S_A indicates agricultural share)

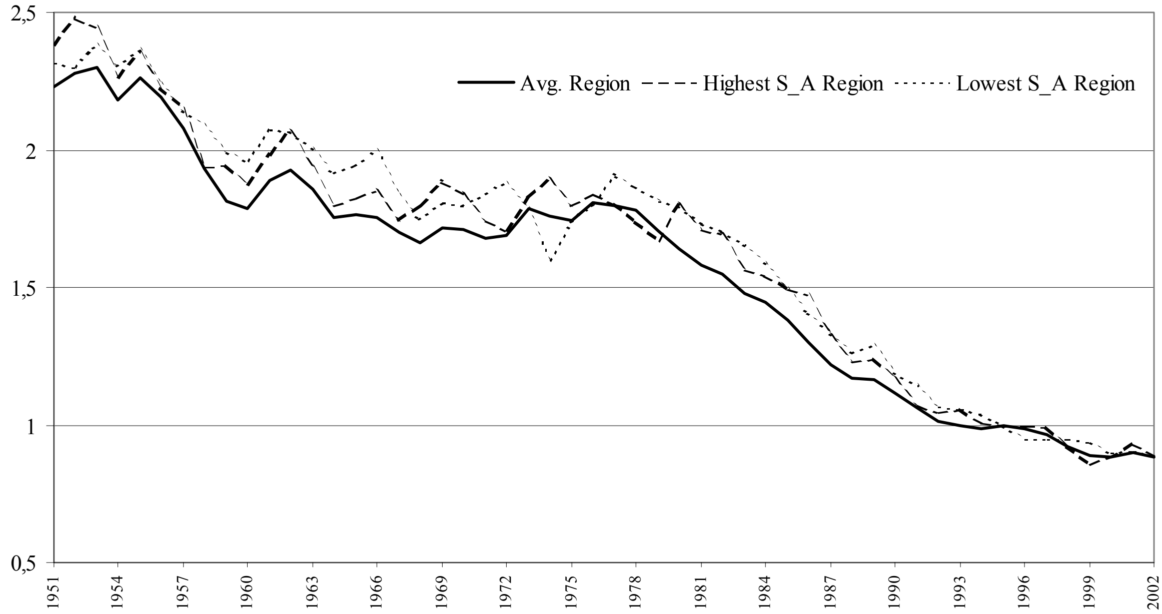


Figure 3 – Evolution of $\mathbf{k}_t$ over the sample period: comparison between the average and the extreme regions (S_A indicates agricultural share)

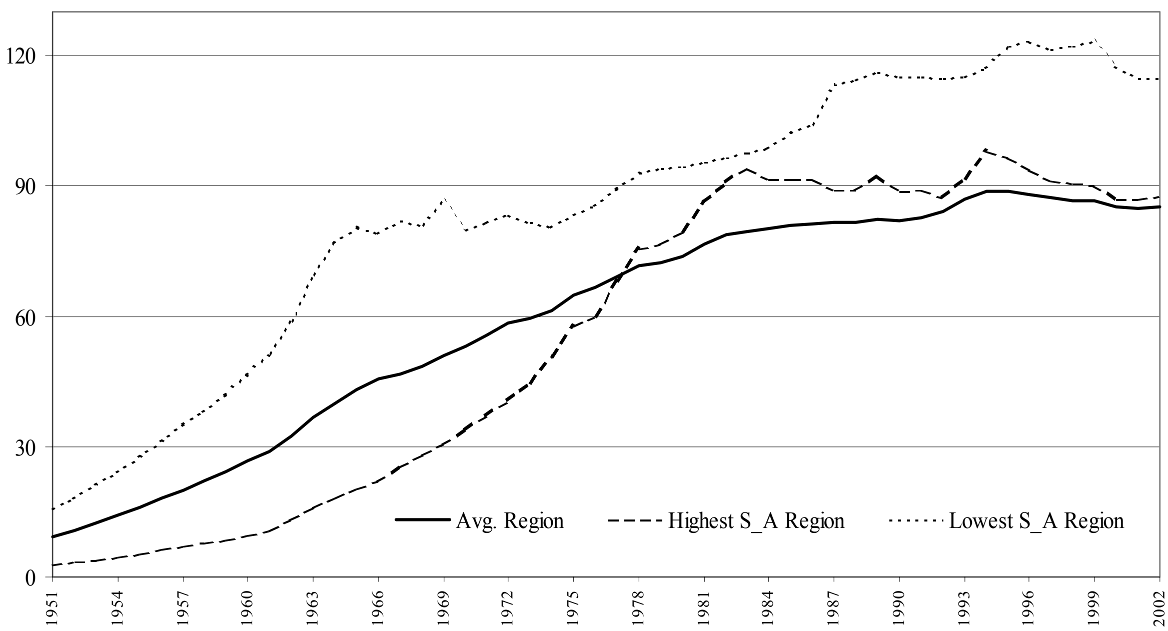


Figure 4 – *Evolution of* $\mathbf{S}_{At}$ *over the sample period: comparison between the average and the extreme regions (S_A indicates agricultural share)*

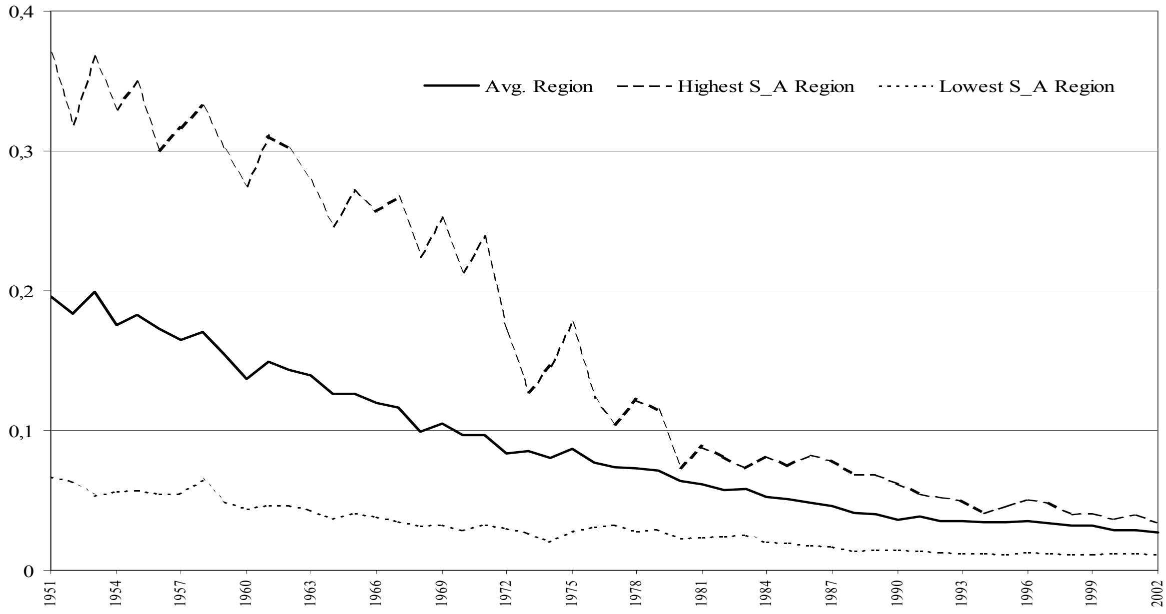


Figure 5 – Evolution of $\mathbf{y}_{At}$ *over the sample period: comparison between the average and the extreme regions (S_A indicates agricultural share)*

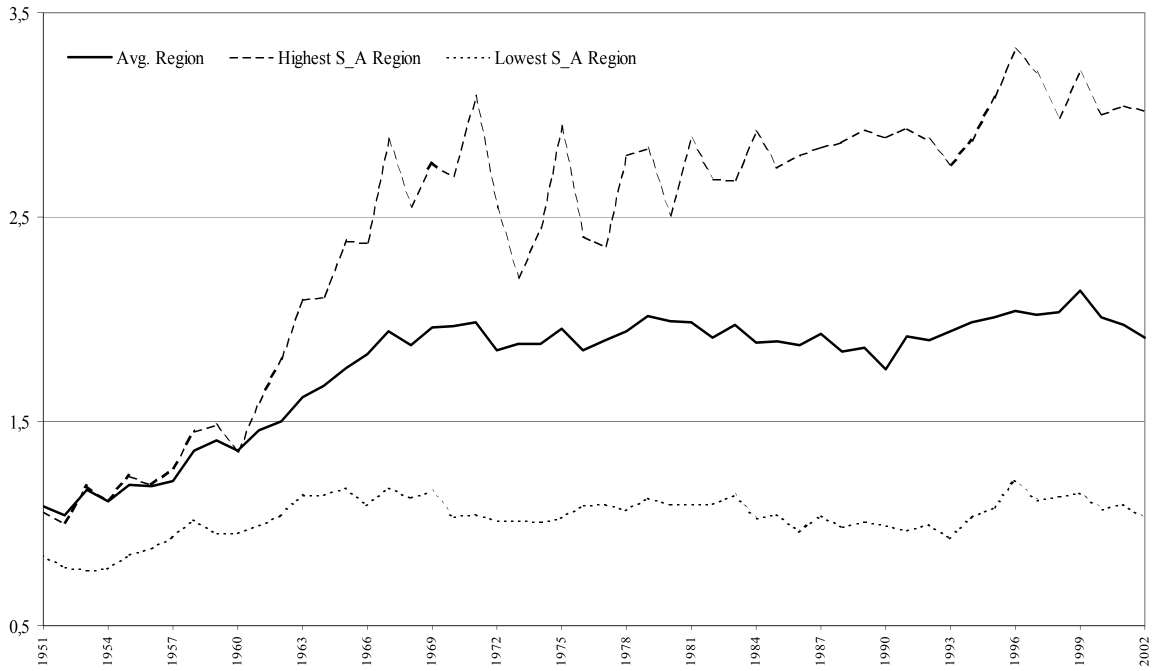


Figure 6 – Evolution of $\mathbf{R}_{At}$ *over the sample period*

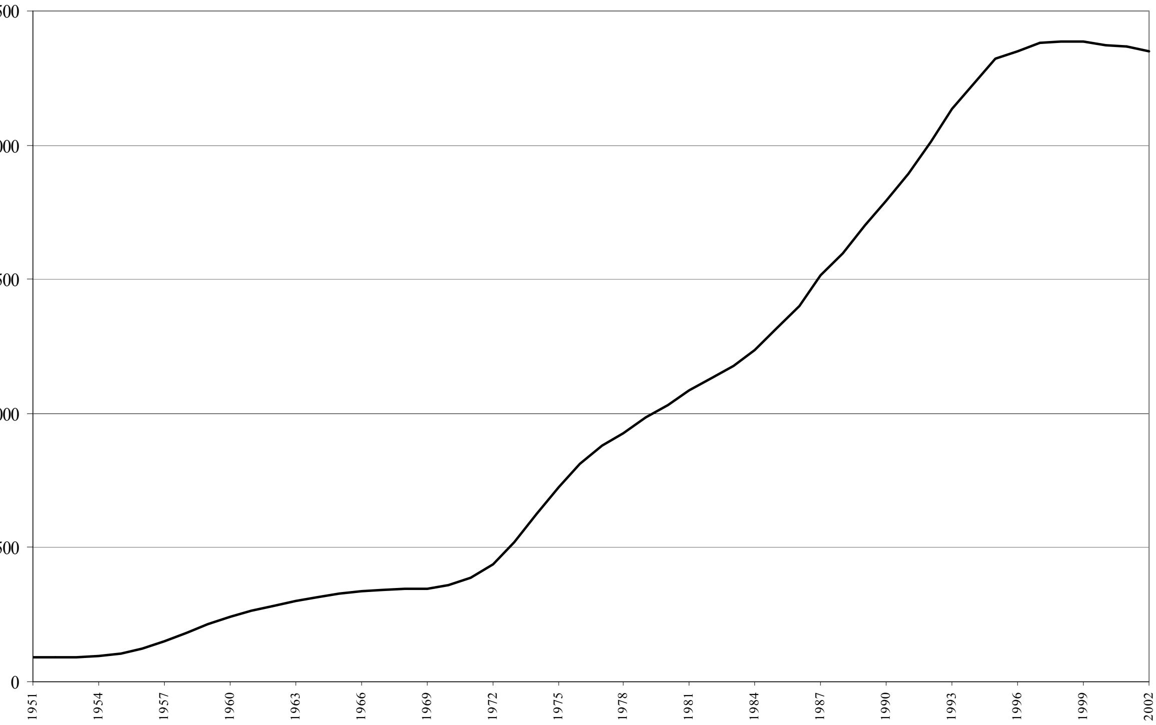
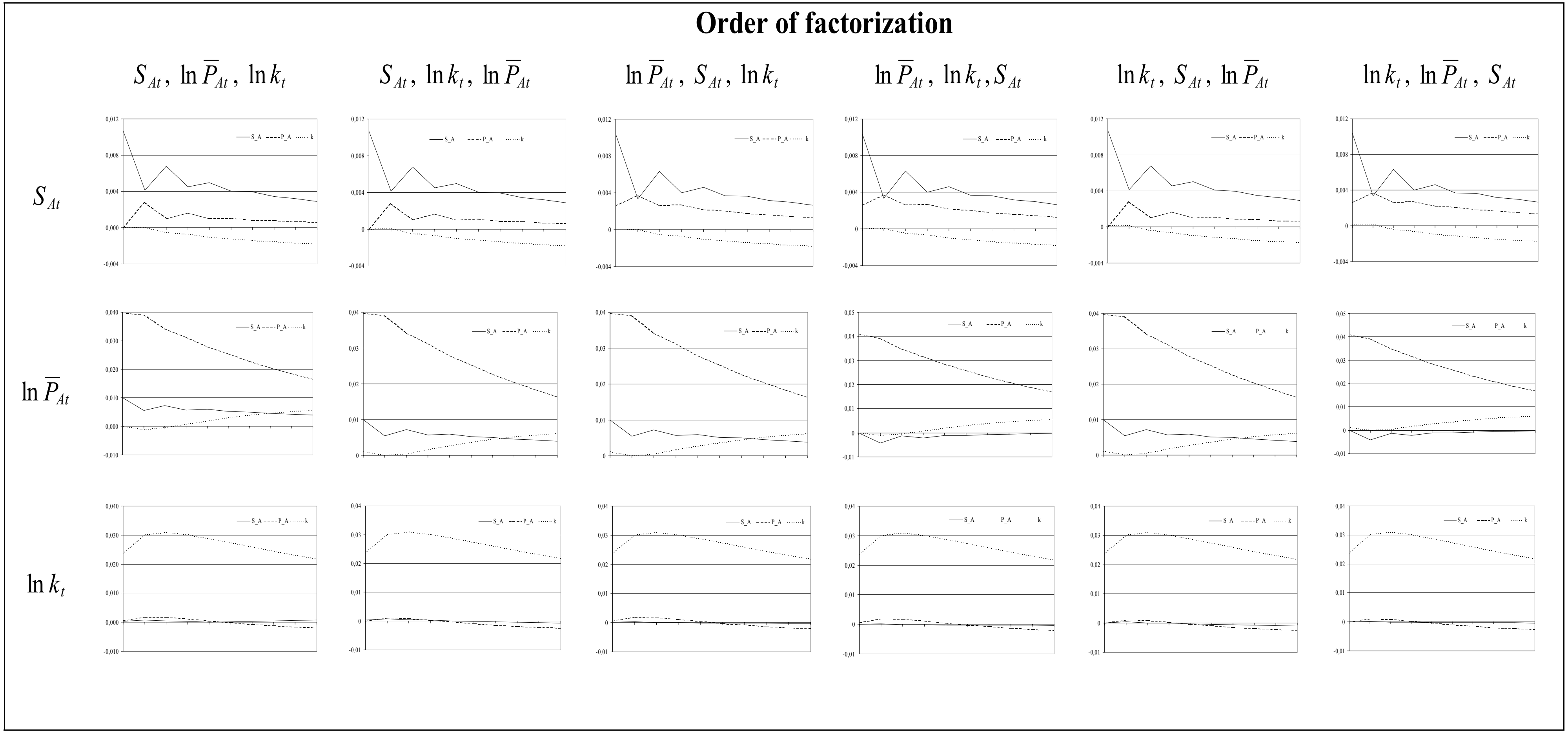


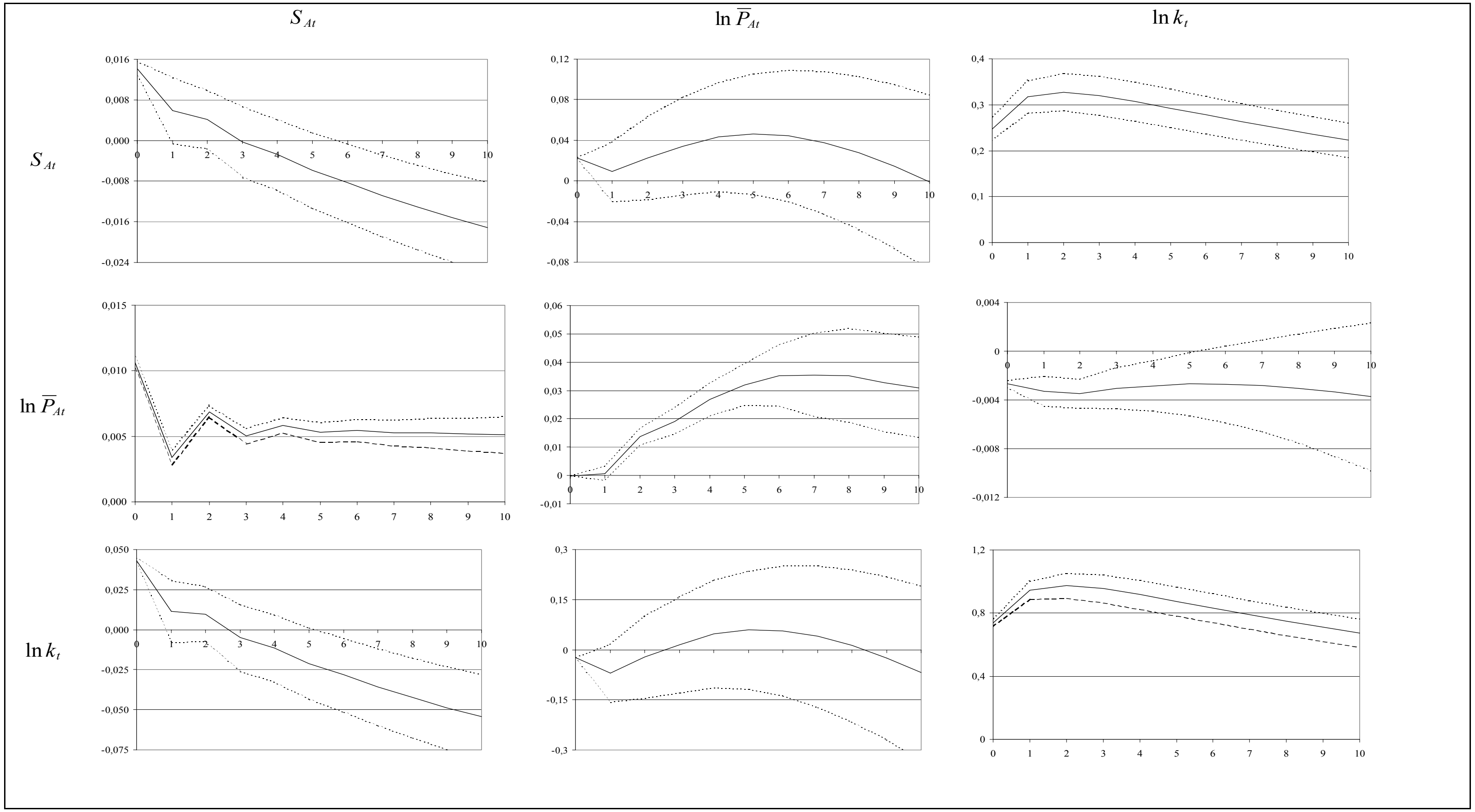
Figure 7 – Recursive VAR: IRFs of endogenous variables (in rows) according to different orderings of factorization (in columns)



Note: Based on the LSDV estimates of the unrestricted VAR (model 1)
Software: TSP4.5

Figure 8 – Structural VAR: IRFs of the endogenous variables (in columns) to respective shocks (in rows)

(dashed lines indicate the 95% confidence interval)



Note: Based on estimates presented in Table 10 (model 4)
Software: STATA9

APPENDIX 1

Derivation of Income Elasticity

Demand income elasticity implied by utility function (16) can be derived by specifying the respective direct demand function, $X_{A,E} = f(\mathbf{P}, I)$, where I=**P′Y** is the disposable nominal income. These demand functions are derived in the two-good case by Primont and Primont (1995):

(A.1) $$X_A = \left(\frac{P_A}{\delta}\right)^{-1/\rho+1}\left[\left(\frac{P_A^{-\rho}}{\delta}\right)^{-1/\rho+1} + \left(\frac{P_E^{-\rho}}{1-\delta}\right)^{-1/\rho+1}\right]^{-1}\left(I - P_A\gamma_A - P_E\gamma_E\right) + \gamma_A$$

(A.2) $$X_E = \left(\frac{P_E}{1-\delta}\right)^{-1/\rho+1}\left[\left(\frac{P_E^{-\rho}}{1-\delta}\right)^{-1/\rho+1} + \left(\frac{P_A^{-\rho}}{\delta}\right)^{-1/\rho+1}\right]^{-1}\left(I - P_A\gamma_A - P_E\gamma_E\right) + \gamma_E$$

where δ and ρ are parameters of the utility function (16). If we assume $\gamma_E = 0$, as in section 4.1.2, we can directly derive income elasticity from the equations above as follows:

(A.3) $$\varepsilon_{X_A,I} = \frac{d\ln X_A}{d\ln I} = \frac{\Lambda I}{\Lambda I - \Lambda P_A\gamma_A + \gamma_A}$$

(A.4) $$\varepsilon_{X_E,I} = \frac{d\ln X_E}{d\ln I} = \frac{\Lambda I}{\Lambda\left(I - P_A\gamma_A\right)}$$

where $\Lambda = \left(\frac{P_A}{\delta}\right)^{-1/\rho+1}\left[\left(\frac{P_A^{-\rho}}{\delta}\right)^{-1/\rho+1} + \left(\frac{P_E^{-\rho}}{1-\delta}\right)^{-1/\rho+1}\right]^{-1}$. Therefore, $\varepsilon_{X_A,I} < 1$ whenever $-\Lambda P_A\gamma_A + \gamma_A > 0$, therefore when $\Lambda P_A < 1$. It follows that $\varepsilon_{X_A,I} < 1$ whenever $\dfrac{\left(\dfrac{P_A^{\rho/\rho+1}}{\delta^{-1/\rho+1}}\right)}{\left[\left(\dfrac{P_A^{\rho/\rho+1}}{\delta^{-1/\rho+1}}\right) + \left(\dfrac{P_E^{-\rho}}{1-\delta}\right)^{-1/\rho+1}\right]} < 1$, and given that *(1-δ)>0*, the inequality always holds true.

On the contrary, $\varepsilon_{X_E,I}$ is clearly always > 0.

Thus, according to how preferences are here modelled, the implicit assumption is that good E is a luxury (or superior) good, while good A always behaves as a normal but non-superior good, regardless the value of the utility function parameters (δ and ρ). These latter parameters only influence the magnitude of income elasticities and, being δ_i a region specific parameter (see section 4.2), income elasticity may differ across regions for both goods, but remains < and > 1, respectively, in all observed regions and years.

APPENDIX 2

Approximation of Equation (17e)

To make the model described in the main text empirically workable, it is particularly useful to approximate $\ln\frac{S_A}{(1-S_A)}$ with $f(S_A)$, where *f(.)* is a polynomial of order *n*, that is $f(S_A)=\sum_{n=0}^{N} c_n S_A^n$, and c_n is a constant term. Given that $0 < S_A < 1$, this approximation could be obtained as Taylor expansion of function $\ln\frac{S_A}{(1-S_A)}$ around a value of S_A within the [0,1] interval. Actually, in the sample here used, S_A always lays in the [.01,.36] interval (see Table 1). In this case, it is easier to pursue the best approximation $f(S_A)=\sum_{n=0}^{N} c_n S_A^N$ just looking for the OLS c_n coefficients over this interval, and stopping at that n-th term when the respective coefficient becomes statistically not significant.

It follows that the third order polynomial (i.e., n = 3) is an acceptable approximation: $\ln\frac{S_{Ai}}{(1-S_{Ai})}=c_0+c_1 S_{Ai}+c_2 S_{Ai}^2+c_3 S_{Ai}^3+\widetilde{u}_i$, where $c_0=-4.5$, $c_1=32.8$, $c_2=-116.3$, $c_3=154.9$, and $\widetilde{u}_i$ is a region-specific error term generated as rest of this approximation.[78]

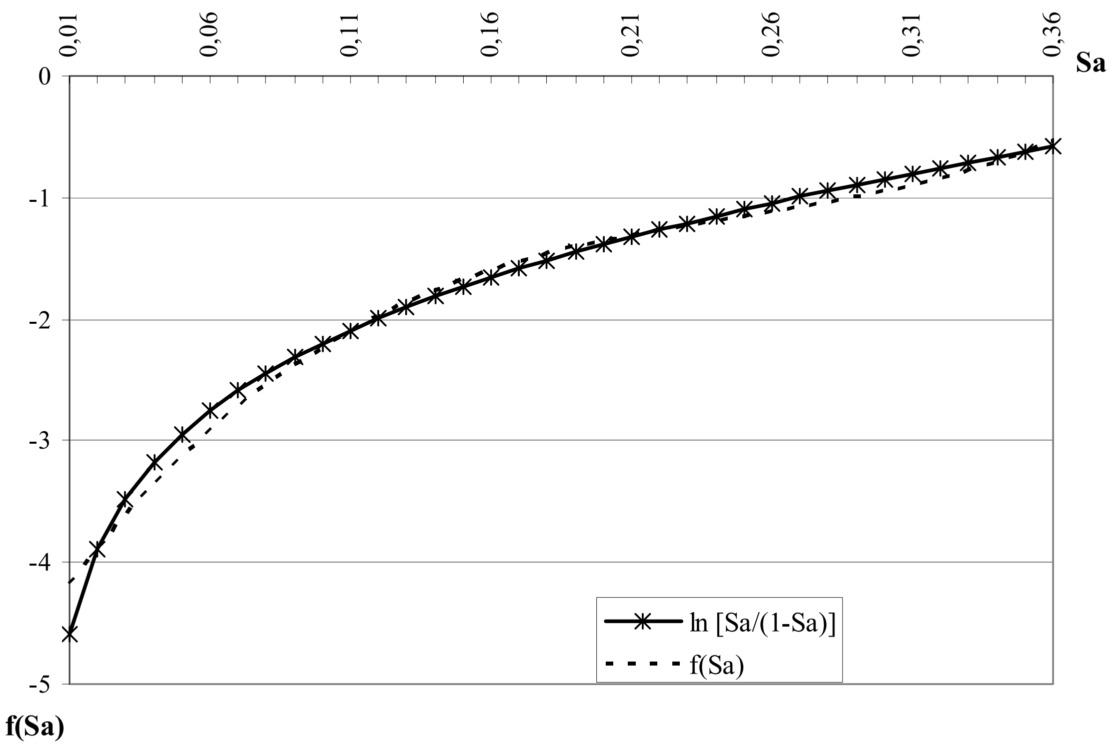
The maximum value of $\widetilde{u}_i$ occurs for $S_A=.31$ and is $u_i=-.09$, 10% of the value to be approximated; in general terms the error is on average very low as indicated by the R^2=.989 of this OLS approximation, and clearly shown in Figure A.1. Consider also that more than 90% of observations fall in the [0.02, 0.25] interval where errors are much smaller, always lower than 5%.

However, $\widetilde{u}_i$ does not only vary across regions but also, for the same region, over time as it depends on S_{Ait}. Consequently, we can write: $E(\widetilde{u}_{it})=0$, $Var(\widetilde{u}_{it})=\sigma^2$, $\forall i=1,...,N$ and $\forall t=1,...,T$; $\widetilde{u}_{it}$ thus behaves as a conventional i.i.d. disturbance term.

---

78 Though this approximation still implies non linearities in one endogenous variable within the model (i.e., S_A) (see section 5.2), it is much more empirically tractable than $\ln\frac{S_A}{(1-S_A)}$, as discussed in section 4.

Figure A.1 – Approximation of function $\ln \frac{S_A}{(1-S_A)}$ *with the 3rd order polynomial function,* $f(S_A) = -4.5 + 32.8\,S_A - 116.3\,S_A^2 + 154.9\,S_A^3$



APPENDIX 3

Construction of Variables

This appendix details data sources, as well as definition and further elaborations of model variables. For all variables the coverage is 20 regions and 52 years (from 1951 to 2002).
Endogenous variables:
- **Agricultural share on regional GDP**$(\mathbf{S}_{\mathbf{A}t})$: it is the ratio between the regional agricultural GDP and the regional GDP both expressed in nominal terms (€). The agricultural VA regional series are taken from the 1951-2002 AGREFIT database (Rizzi and Pierani, 2006). The regional GDP series are taken from the CRENoS Regio-It 1951-93 and Regio-It 1970-04 harmonized databases (Paci and Saba, 1997).
- **Agricultural relative price**$(\overline{\mathbf{P}}_{\mathbf{A}t})$: it is the ratio between the agricultural implicit price index and the aggregate implicit price index (both in 1995 base). The former is the deflator of the value of regional agricultural production (expressed in nominal €) and is taken from the 1951-2002 AGREFIT database. The latter is the deflator of the regional non agricultural-GDP (expressed in nominal €) and is taken from the CRENoS Regio-It 1951-93 and Regio-It 1970-04 harmonized databases.
- **Capital *per* labour unit or capital intensity**$(\mathbf{k}_t)$: it is the ratio between the amount of regional stock of physical capital (expressed in millions € at 1995 constant prices) and the total regional labour force. The regional capital stock is computed from the gross total investment series (expressed in millions € at 1995 constant prices) as reported by the CRENoS Regio-It 1951-93 and Regio-It 1970-04 harmonized databases. The 1951-2002 gross investment series have been transformed into the capital stock series by, firstly, prolonging backward the series according to the methodology proposed by Ball *et al.* (1993), and, then, by applying the method of Hulten (1990; see also Diewert, 2005) for the calculation of the productive stock. Taking the average values reported by Paci and Pusceddu (1999) and Lupi and Mantegazza (1994) we assume: an average life of 18 years but possibly ranging between 1 and 36 years over a truncated normal distribution of withdrawals; a real long-term interest rate of 4%; a convex age/efficiency curve, expressing the declining efficiency of the investment over time, with parameter β=0.5 (see also Rizzi and Pierani, 2006, for details). The series of the regional work force (expressed in thousand units) are taken from the CRENoS Regio-It 1951-93 and Regio-It 1970-04 harmonized databases. k_t is thus expressed in 1995 thousands € of capital stock *per* single labour unit.
Exogenous variables:
- **Agricultural production *per* labor unit** $(\mathbf{y}_{At})$: it is the ratio between the regional agricultural production (expressed in millions € at 1995 constant prices) and the regional labour force (expressed in thousand units). The variable is thus expressed in 1995 thousands € of agricultural production *per* labour unit. The regional series of agricultural production are taken from the 1951-2002 AGREFIT database. The series of regional work force (expressed in thousand units) are taken from the CRENoS Regio-It 1951-93 and Regio-It 1970-04 harmonized databases.

- **Public National Agricultural R&D Stock** $\left(\mathbf{R}_{At}\right)$ **:** it is the amount of national public agricultural R&D stock expressed in 1995 millions €. This stock is computed form the respective public agricultural R&D investment series updated to 2002 and detailed in Esposti and Pierani (2000) and Esposti and Pierani (2006). Investments in nominal terms have been deflated according to the specific public agricultural R&D deflator calculated by Esposti and Pierani (2006). The R&D stock series have been then computed from these investment data using the methodology and parameters reported in Esposti and Pierani (2003). Following this approach, the stock series thus cover the whole 1951-2002 period.
- **Regional Dummies** $\left(\mathbf{D}\right)$: N-1 regional dummy variables taking value 1 for the respective region and 0 for all the other cases. To avoid singularity the dummy of Valle d'Aosta region is skipped.
- **Trend** $\left(\mathbf{t}\right)$: it is the conventional trend variable going from 1 (in 1951) to 52 (in 2002) over the whole sample period.
- **CAP1:** it is a time dummy variable aimed at seizing the effect of the introduction of the CAP, especially price support system; thus, it takes value 0 from 1951 to 1962 and 1 from 1963 onwards.
- **CAP2:** it is a time dummy variable aimed at capturing the effect of the new CAP regime introduced in 1992 by the so-called MacSharry Reform; thus, it takes value 0 from 1951 to 1992 and 1 from 1993 onwards.

APPENDIX 4

Table A.1 – SVAR model estimates of region-specific effects

Equation:	$\ln\overline{\mathbf{P}}_{At}$		$\ln\mathbf{k}_t$		$\mathbf{S}_{At}$	
	Parameter Estimate	*Standard Error*	*Parameter Estimate*	*Standard Error*	*Parameter Estimate*	*Standard Error*
Constant	-.0225	.0353	.1060*	.0241	.0347*	.0091
<u>Northern</u>						
Friuli	-.0055	.0113	-.0112*	.0049	-.0124*	.0025
Liguria	-.0183	.0097	-.0141*	.0048	.0023	.0022
Lombardia	-.0044	.0099	-.0128*	.0048	-.0073*	.0023
Piemonte	-.0073	.0119	-.0115*	.0050	-.0139*	.0026
Trentino Alto-Adige	-.0072	.0119	-.0064	.0049	-.0128*	.0026
Veneto	-.0038	.0142	-.0128*	.0051	-.0186*	.0032
<u>Central</u>						
Abruzzo	.0031	.0149	-.0101	.0053	-.0166*	.0033
Emilia-Romagna	-.0013	.0167	-.0133*	.0054	-.0223*	.0037
Lazio	-.0080	.0100	-.0163*	.0049	-.0074*	.0023
Marche	.0001	.0144	-.0161*	.0053	-.0205*	.0032
Toscana	-.0115	.0109	-.0175*	.0051	-.0103*	.0024
Umbria	-.0091	.0132	-.0097	.0050	-.0133*	.0030
<u>Southern</u>						
Basilicata	-.0069	.0149	.0008	.0060	-.0061	.0034
Campania	-.0155	.0123	-.0119*	.0051	-.0098*	.0027
Calabria	-.0134	.0137	-.0088	.0053	-.0114*	.0030
Molise	.0086	.0169	-.0043	.0059	-.0189*	.0038
Puglia	-.0125	.0157	-.0181*	.0056	-.0167*	.0034
Sardegna	-.0115	.0136	-.0052	.0052	-.0112*	.0030
Sicilia	-.0098	.0145	-.0096	.0054	-.0156*	.0032

Note: Results refer to SVAR model estimate reported in Tables 10 and 11; to avoid singularity, the dummy variable of region Valle d'Aosta was excluded
** Statistically significant at the 5% level*

Table A.2 – Decomposition of regional agricultural share variation (1951-2002)

	Total ΔS_{At}	Own effect $\theta_{S_A}\ \Delta S_{At}$	Components (equation (25))		
			Relative price effect $\theta_{P_A}\ \Delta \ln \overline{P}_{At}$	Capital intensity effect $\theta_k\ \Delta \ln k_{it}$	Technological gap effect $\phi_R \Delta \ln R_{At} + \phi_t\ t$
<u>*Northern*</u>					
Friuli	-.0795	.0039	14.6%	113.6%	-28.2%
Liguria	-.0629	.0031	22.9%	13.6%	-53.5%
Lombardia	-.0860	.0042	24.2%	117.2%	-41.4%
Piemonte	-.1059	.0052	15.3%	116.3%	-31.6%
Trentino Alto-Adige	-.0883	.0044	12.4%	114.3%	-26.7%
Veneto	-.1797	.0089	15.6%	114.9%	-3.5%
Valle d'Aosta	-.0561	.0028	19.5%	117.5%	-37.0%
<u>*Central*</u>					
Abruzzo	-.2158	.0106	12.6%	113.5%	-26.1%
Emilia-Romagna	-.2761	.0136	16.5%	114.3%	-3.8%
Lazio	-.0874	.0043	19.8%	12.7%	-4.5%
Marche	-.2056	.0101	16.9%	109.4%	-26.3%
Toscana	-.1025	.0051	14.7%	122.9%	-37.6%
Umbria	-.0561	.0028	19.8%	114.0%	-33.8%
<u>*Southern*</u>					
Basilicata	-.2834	.0140	2.7%	113.9%	-34.6%
Campania	-.1496	.0074	14.3%	119.0%	-33.4%
Calabria	-.1851	.0091	13.5%	113.8%	-27.3%
Molise	-.3355	.0165	1.9%	109.1%	-2.0%
Puglia	-.2536	.0125	13.7%	119.0%	-32.7%
Sardegna	-.1746	.0086	14.9%	121.3%	-36.2%
Sicilia	-.2836	.0140	14.2%	113.3%	-27.5%

Map A.1 – Regional change of S_{At} (A) and % capital intensity effect (B) (1951-2002)