Stichting Onderzoek Wereldvoedselvoorziening van de Vrije Universiteit

Centre for World Food Studies

The Chinagro model:
a spatially detailed general equilibrium welfare model of China's agricultural economy

by

Michiel A. Keyzer and Wim C.M. van Veen

Staff Working Paper **WP - 12 - 01** **March 2012**

Contents

Abstract

The paper gives a mathematical description of the Chinagro model, version 2.0, a spatially detailed general equilibrium welfare model of the Chinese economy with emphasis on the agricultural sector. It has been developed by SOW-VU in cooperation with the Center for Chinese Agricultural Policy (CCAP) in Beijing and the International Institute of Applied Systems Analysis (IIASA) in Laxenburg, Austria. The model has been used for scenario simulations of the future of China's agricultural economy in two successive EU-funded projects, viz. CHINAGRO (2001-2005) and CATSEI (2007-2010), and is updated regularly.

To account for the geographical differences in terms of resources and output potentials, the model distinguishes farms at county level, about 2800 in number. Every farm maximizes its net revenue at given output and input prices. Available land and labor resources depend on autonomous drivers as well as on economic growth outside agriculture, and rural-to urban migration. On each farm, land uses in cropping and animal husbandry compete for available labor and machinery. They also exchange local commodities such as organic manure and crop residuals. For each land use type technology is represented by production functions with two branches, a piecewise linear branch that indicates how much fertilizer or feed is required to achieve a given yield, and a Mitscherlich-Baule branch that does the same for labor. The program explicitly maintains material balances of nutrients, land, labor, livestock numbers and stable capacity. Optimal farm response leads to outputs and inputs that respond continuously to prices.

Demand is represented at more aggregate level, viz. by income group and location (urban/rural) in eight regions, via utility maximizing consumers who maximize utility according to a linear food expenditure system with time-dependent coefficients and quasi-linearity in non-food. Traders minimize the costs of trade and transportation of sixteen agricultural commodities within and across the eight regional markets. Government tax and trade policies are imposed exogenously. Foreign preferences for net imports from China are represented via a trade welfare function that is maximized in competition with the welfare of Chinese consumers, thus covering the impact of China's imports and exports on world prices.

The model is recursively dynamic and applied to a selected number of years over the period 2005-2030. Calculations proceed in GAMS. In each year the equilibrium solution is obtained by combining use of the GAMS solver with a dedicated algorithm for agricultural supply and price feedback rules. Base year outcomes are fully calibrated to replicate the model's independently constructed internally consistent 2005 data set.

1. Introduction

Chinagro is the simulation model that has served as backbone for two successive EU-funded projects, CHINAGRO (2001-2005) and CATSEI (2007-2010).[1] Both projects have studied China's prospects in food and agriculture in a rapidly changing domestic and international setting, and the impacts on foreign trade, social conditions in rural areas and environmental pressure. In both projects, the Chinagro model has served to integrate findings from separate detailed studies in these domains within a coherent framework.

Chinagro 1.0 is the model version developed in the first project. Its database focuses on the base year 1997, with validation over the period 1997-2005, and simulations until 2030. The CATSEI project extended it into version 2.0, shifting the base year to 2005, with validation until 2010, and updating the geographical classification to reflect new administrative subdivisions into counties. Version 2.0 also has an enhanced and more detailed specification of agricultural supply, and rather than treating world market prices as given, accounts for the impact of China on international markets.

The present paper describes the structure of Chinagro 2.0, henceforth referred to as Chinagro for short. A selection of scenario outcomes is discussed in a companion paper (Keyzer et al., 2012). This model structure is designed to account for the fact that China's agriculture cannot be looked at in isolation. Its relations to other sectors, and to other countries are key drivers. Similarly, within rural areas, agriculture is more than farming alone. There are interactions between crop and livestock subsectors, there is competition for inputs among these, and agricultural transition goes hand in hand with demographic and technological change that all impact on the environment. More specifically, Chinagro is designed to accommodate the following characteristics of China's agricultural economy in transition.

First, the mere size of the country, its wide range of climatic, soil and other natural conditions, with different vulnerability to environmental degradation, the uneven spread of population across the country, and the long distances between population centers call for a spatially explicit model.

Second, given China's economic size and in view of its present orientation on foreign trade, proper attention should be paid to China's links with and impacts on world agricultural markets, on the import side particularly regarding the feed grains needed for China's livestock sector and for biofuels, and on the export side the potential outlets for Chinese fruits and vegetables. Indeed, size can become a handicap, as fast rise in imports will tend to raise world prices and conversely for exports. This has to be reflected, with special reference to the policy measures taken in the context of China's accession to the World Trade Organization (WTO) in 2001 and as reaction to price hikes on the world markets in recent years.

Third, the model should incorporate various linkages with sectors outside agriculture, through processing and transport, through competition for land and through migration from rural to urban areas, leading also to workers' remittances to rural households.

[1] Cooperating partners in the CHINAGRO project were the Centre for World Food Studies, Amsterdam (SOW-VU), the Center for Chinese Agricultural Policy, Beijing (CCAP) and the International Institute for Applied Systems Analysis, Laxenburg (IIASA), which acted as coordinator. The same partners cooperated in the CATSEI project, this time with SOW-VU as coordinator, now joined by the School of Oriental and African Studies, London (SOAS), the Agricultural Economics Research Institute, The Hague (LEI), and the International Food Policy Research Institute, Washington DC (IFPRI).

Finally, shifts in consumption patterns should be accounted for with increasing incomes and rural-to-urban migration itself, with rising expenditure shares for meat and dairy and for fruits and vegetables, and declining shares for staple foods such as rice, wheat and tubers.

The project has opted for a general equilibrium approach to address these issues, specifically a spatially explicit welfare program. The model is recursively dynamic, and applied to a selected number of years over the period 2005-2030 following the specification of carefully designed scenarios. Calculations proceed in GAMS (Brooke et al., 2011). In each year the equilibrium solution is obtained by combining use of the GAMS solver with a dedicated algorithm for agricultural supply and price feedback rules. Base year outcomes are fully calibrated to replicate the model's independently constructed internally consistent 2005 data set. Outcomes are presented via provincial, regional and national tables and via maps, the latter especially for results at county level.

Chinagro falls with the tradition of general equilibrium theory (Debreu, 1959; Arrow and Hahn, 1971), in the class of computable models generally referred to as Applied General Equilibrium or AGE for short (Gunning and Keyzer, 1995; Ginsburgh and Keyzer, 2002). When looked at as a world model it differs from the international model of the Global Trade Analysis Project (GTAP, Hertel, 1997) that also comprises a China module and adopts the Computable General Equilibrium (CGE) approach as pioneered by Dixon et al. (1977) and Dervis et al. (1982), in four respects: (i) its representation of the rest of the world is rudimentary, (ii) its principle of commodity classification does not per se treat foreign commodities as different from domestic ones, (iii) the formulation of agricultural supply is dedicated to China's specificities instead of following a standard functional form without reference to resource availability, and (iv) unlike in common CGE-models material balances are being respected across commodity accounts, so that, for instance, the final product can never contain more calories or minerals than the primary product.

Chinagro can also be compared to partial equilibrium models of China's agriculture, several of which are being used as building blocks of worldwide agricultural modeling systems. We mention the AGLINK-COSIMO model of the Food and Agricultural Organization of the United Nations and the Organisation for Economic Cooperation and Development (OECD, 2007), the world model of the US-based Food and Agricultural Policy Research Institute (FAPRI, 2011) and the IMPACT model of the International Food Policy Research Institute (Rosegrant et al., 2008). These models have econometrically estimated, price-responsive supply and demand equations for all agricultural commodities, distinguishing on the supply side harvested area and herd size as well as yield, while simultaneous market clearing leads to world and country prices. Of the agricultural simulation models used in China itself, the most frequently cited one is CAPSiM, a partial equilibrium model maintained at CCAP (Huang and Li, 2003). CAPSiM clears agricultural markets at national level but its recent versions further distribute supply outcomes to provinces and food intake to rural household income groups.

Chinagro differs from these partial models essentially in three respects, both concerning the supply formulation: (i) it relies on the interplay between structure, classification, resource balances, agronomic information and estimation more than on direct econometric estimation; (ii) it has far more spatial detail, taking into account even local non-traded goods, whereas the partial models operate at the national level; and (iii) it is a general equilibrium model, but this element of distinction may be less important as Chinagro's non-agricultural supply is exogenous anyhow and its utility functions are quasi-linear with consumption of non-agricultural product as numeraire good, which makes consumption of agricultural products insensitive to endogenous changes in income.

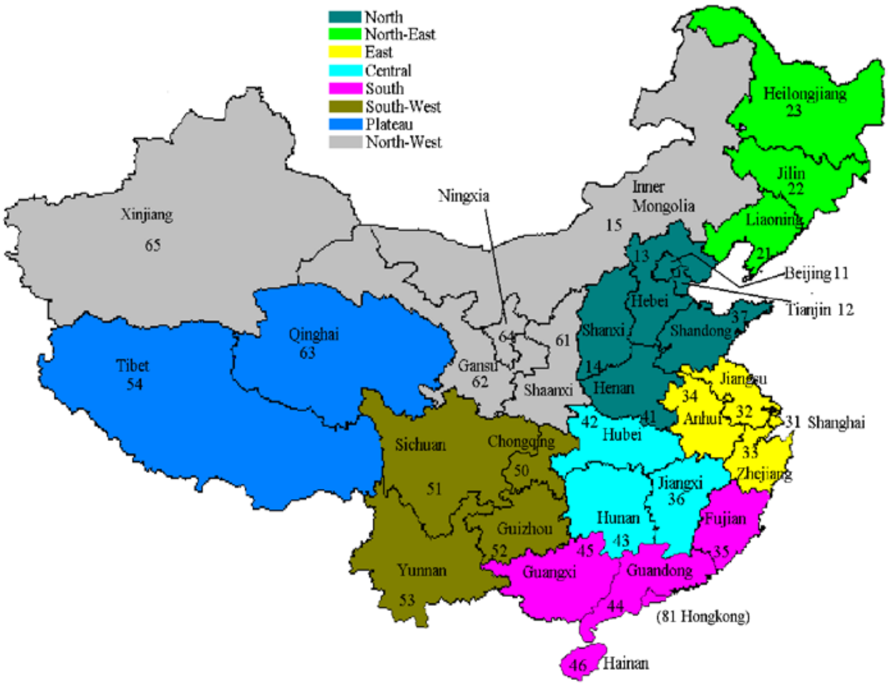
The paper proceeds as follows. First, section 2 gives an overview of the model, in a non-technical way. Then, the welfare model is formally introduced in section 3, in two steps (simple and full version), with agricultural supply still generally formulated, like a black box, that we open in section 4 where we detail the agricultural supply component that runs by county. Section 5 briefly discusses some practicalities of model development and application. Model classifications appear in the Appendix.

2. A bird's eye view of the model

The general equilibrium model used for scenario simulations is cast in the form of a single-period welfare program that is solved for selected years of simulation over the period 2005-2030, evaluating solutions under given scenario conditions with respect to resource availability, demography, non-agricultural growth, life-style changes, technological progress, international prices and government policies. The years selected for simulation are 2005, 2010, 2020 and 2030. With respect to validation, the model fully replicates for every county and region of China in the 2005 base-year conditions, adequately mimics changes over the period 2005-2010 and provides interpretable results until 2030.

The welfare program has eight regions, with farm supply described at the much more disaggregated county level. The regions are shown in Figure 2.1. The distinction between the regions is based on their respective geographic, agro-climatic and demographic features, and economic development levels. The regions are subdivided into provinces, the actual administrative units in China.

Figure 2.1 Map of China with provincial boundaries and the eight Chinagro regions



Note: Taiwan is not included in this analysis and, therefore, absent from the map.

Farm supply is modeled for each of 2885 counties, covering the whole of China, with farm output cast in terms of various activities covering fourteen types of crops and nine types of animals, whereas also related activities such as collection of household waste and supply of machinery power are included. Crop activities take place on three land use types (irrigated land, rainfed land, orchards), whereas the livestock activities are conducted in six different livestock systems, distinguished on the basis of mode and intensity of production. The output of the cropping land use types and livestock systems comprises both local and tradable commodities. Local commodities are traded only inside the own county, and even then only over limited distances. They consist of local feed (grass, crop residuals, household waste, green fodder), organic fertilizer (animal manure, nightsoil) and power (draught power, machinery power). Tradable commodities are exchanged across all regions and from and to abroad. Their prices are determined endogenously in the general equilibrium model. The model distinguishes seventeen of such tradable commodities, of which thirteen are food commodities and two are feed commodities, while one commodity (maize) is used for both food and feed and one commodity is non-agricultural. Appendix A presents the supply and commodity classifications.

The general equilibrium welfare model focuses on the description of (a) supply response by farmers under their prevailing technology and natural resource endowments by county, (b) the behavior of consumers by region and income group for rural and urban separately, (c) the balancing on the regional markets of supply and demand, with appropriate trade between regions and with the foreign markets. Together, these decisions of producers, consumers and traders are such that at given exogenous conditions (in particular, agricultural resources and technology, non-agricultural output, world price trends, government policies and the country's trade surplus) and at given welfare weights, optimal social welfare is obtained. Indeed, once market distortions have been eliminated the model in every particular year generates an optimal allocation of agricultural production among regions, based on comparative advantage, while accounting for transportation costs.

Non-agricultural supply and demand is specified by region (hence, not at county level) and largely exogenously, thus setting the level of overall economic activity the agricultural sector operates in. Also output from fishery and forestry are represented exogenously. The domestic non-agricultural price remains fixed and normalizes all other prices, as is necessary for this commodity to act as numeraire. Therefore, all resulting prices and expenditures can be interpreted as "real" and comparable to the 2005 price and expenditure levels.

The specification of the model is best understood in terms of the behavior of the individual agents that appear in it, i.e. consumers, farmers, traders and the rest of the world:

Consumers

Consumers are distinguished by location (urban, rural), by region and by income group (poor, middle, rich). Each individual of a specific group spends revenue on food according to a linear expenditure system with time-dependent coefficients. As mentioned earlier, expenditures on non-food appear as closing item on individual budgets of consumer groups and in the national government budget. Revenue originates from direct earnings (entrepreneurial income, factor rewards) and private transfers as well as from government transfers (which are negative in case they pay income tax). Welfare weights on group utility determine the level of uncommitted expenditures of each consumer group. Hence, consumer demand adjusts both to scenario variables and to variables determined by the model itself, the endogenous variables.

Farmers

Farms are distinguished by county. The typical farm of a county chooses its cropping pattern and livestock activities by allocating its labor and equipment so as to maximize its current revenue, i.e. the net proceeds from sales minus the cost of current inputs (purchased feeds for animals and fertilizer for crops), subject to technological constraints specified separately for cropping land use types (irrigated land, non-irrigated land, orchards) and for several types of livestock systems (ruminant, non-ruminant with varying degree of intensification), for given land areas and stable capacities.

The technology of each land use type (including the livestock systems) is represented via a two-branch production function. One branch indicates how much fertilizer per hectare (respectively, feed per stable unit) is needed to achieve a given yield, the other does the same for labor per hectare (respectively, labor per stable unit). The fertilizer or feed branch is piecewise linear, starting from a positive yield at zero formal input, using only locally available manure and nightsoil (respectively, locally available feed from crop and household residues) and exhibiting positive but decreasing returns from purchased fertilizer (respectively, purchased feed). The labor branch is of the Mitscherlich-Baule type, starting also from a positive yield at zero formal input (using only informal work or child labor), having also positive and decreasing returns to formal input but with the yield not exceeding a bio-physical cropping potential (asymptote) measuring the maximal amount of biomass that could be produced through photosynthesis under the given climate and soil conditions.

Purchased fertilizer and feed determine farm demand for tradable commodities. Local commodities (organic fertilizer and local feeds) reduce this demand by their impact on the yield level at zero formal inputs. Labor input is expressed in terms of equipped labor equivalents, comprising labor as well as available animal draught power and machinery. In each county, available (equipped) farm labor is optimally allocated across land use types.

On the output side, every land use type produces several commodities, according to substitution possibilities among different crop and livestock activities within each land use type, while maintaining area and stable capacity balances. For example, non-ruminant farm types jointly produce pork, poultry and eggs in county-specific proportions that can change under shifts in the relative prices of these goods. In addition, this farm type produces manure as a byproduct that can be used as fertilizer, and hence substitutes for purchased fertilizer. Similarly, cropping systems produce various tradable goods such as grains, vegetables, fruits and marketed feeds, as well as crop residuals such as straw and husks that can be used as local source of feed for livestock. Clearly, ruminants can use feed from pastures and grazings as well as other types of roughage, while non-ruminants are more restricted in their diet. Exogenous, county-specific utilization rates account for the limited possibility to use all byproducts from farming.

Traders

Traders minimize for every traded commodity the total cost of delivery they must incur to satisfy consumption and input demand, given (i) the supply in the various counties, (ii) the possibility to import from and export to the world market at a given, tariff ridden price possibly subject to quota on foreign trade, and (iii) the unit cost of transport between regions as well as the unit cost to process the agricultural product from county level up to consumer level. This leads to trade levels as well as regional and county prices at which deliveries take place. For this trade module, government taxes, tariffs and quotas are the scenario parameters describing the policy being implemented. Furthermore, the foreign price developments and the trend assumptions on the unit cost of trade, transport and processing enter as scenario variables.

Rest of the world

The rest of the world is represented via a foreign consumer whose preference for net import from China is modeled as trade welfare function that is being maximized in competition with Chinese consumers.

3. General equilibrium and the Chinagro-welfare model

3.1 Welfare and competitive equilibrium

We now formulate the applied general equilibrium model. Our key simplifying assumption is that utility functions are quasilinear (e.g. Varian, 1992), which ensures that the equilibrium can be implemented as a convex program. We refer to Chinagro as a welfare model, because it implements this convex program as welfare optimum distorted by prevailing indirect taxes and tariffs.

This quasi-linearity assumption means that we can abstain from imposing a budget constraint to every consumer group separately. We consider this appropriate for the following reasons. First, our description of the non-farm sector in the Chinese economy is rudimentary at best. This makes it difficult to derive a realistic distribution of primary income over household groups. Second, data available on distribution within regions are more reliable for expenditures than for revenue, particularly because of scarce information on transfers from migrant workers. Information on the distribution of savings is not available either. Third, several public redistribution mechanisms are currently in place that would have to be analyzed as well. Finally, while the model's focus is on production and consumption of agricultural commodities, a more explicit representation would be very demanding, because given the size of the non-farm sector, inevitable simplifications will tend to dominate model outcomes.

Chinagro simulations consist of a sequence of static solutions of this welfare model for selected years. Every year describes an equilibrium among a large number of actors, with (i) consumers maximizing their surplus, and (ii) producers and traders their profits, while (iii) competitive markets provide for the linkages between them.

3.2 Specification of the welfare model: basic version

To avoid simultaneous presentation of too many aspects, we introduce the Chinagro model in two versions, a basic and a full one. The basic version of the static welfare model is specified as follows:

(1) We distinguish counties indexed $c,\ c = 1,...,2885$ and aggregate from counties to regions indexed $r,\ r = 1,...,8.$ These indices appear as subscripts and are also taken to distinguish between the region- and county-level for both variables and functions. The set C_r indicates all counties belonging to region *r*.

We also distinguish commodities indexed $k,\ k = 1,...,K$. As a rule, commodity transactions are represented by *K*-dimensional vectors with commodities *k* as elements. Thus, all commodities are considered simultaneously, but some, such as the good used in transport can be purchased at a fixed price, of the numeraire good, taken to be good *K*.

Furthermore, we distinguish *I* consumer groups, indexed *i*, with for every region *r* a subset I_r^u for urban and I_r^r for rural, with three groups each (low, middle and high income); this also defines the set $I_r = I_r^u + I_r^r$. An applied general equilibrium model in welfare format defines welfare as the (weighted) sum of the utility of these consumer groups, all expressed in money metric.

(2) As mentioned above our key simplifying assumption is that we take all these utility functions to be quasilinear. This means that for *K* commodities indexed *k* they are additively separable with respect to the *K*-th (numeraire) good "non-agriculture", and can for every group be written:

$$u_i(x_{i1},...,x_{iK}) = \tilde{\upsilon}_i(x_{i1},...,x_{i,K-1}) + \bar{p}_K x_{iK}, \tag{3.1}$$

where x_{ik} denotes consumption of commodity *k* by consumer *i*, $\bar{p}_K$ is a given price of the numeraire good, and functions $\tilde{\upsilon}_i$ are strictly concave increasing and twice differentiable. For ease of notation we refer in the sequel to utility $\tilde{\upsilon}_i$ in terms of the full consumption vector, by defining $\upsilon_i(x_i) = \tilde{\upsilon}_i(x_{i1},...,x_{i,K-1})$. The Chinagro-model in fact also keeps a component of numeraire consumption in υ_i for normalization purposes but this is immaterial since the remainder of non-agricultural consumption, dealt with as in (3.1), will follow residually at equilibrium.

Similarly, the rest of the world's preferences are expressed via a trade welfare function, $U(z^f)$, which is

$$U(z_1^f,...,z_K^f) = \tilde{U}(z_1^f,...,z_{K-1}^f) + \bar{p}_K z_K^f \tag{3.2}$$

also a strictly concave, twice differentiable utility but with as main difference that net imports z^f are not required to be non-negative, albeit that $\lim_{\min_k z_k^f \to -\infty} U(z^f) = -\infty$, i.e. U is required to become highly negative whenever foreign net supply of any commodity becomes infinite, so as to reflect boundedness of foreign supply.

The corresponding welfare criterion is based on summation of these money metric utilities and a public demand w_0 of numeraire good, for investment as well as consumption. It reads:

$$W = \sum_r \left(\sum_{i \in I_r^u} u_i(x_i) + \sum_{i \in I_r^r} u_i(x_i) \right) + \bar{p}_K w_0 + U(z^f) \tag{3.3}$$

which in a welfare optimum will be seen to imply for all individual consumers that they maximize their surplus:

$$x_i = arg\,max_{x \geq 0}\ u_i(x) - p_i^T x \tag{3.4}$$

and similarly for the foreign sector:

$$z^f = arg\,max_{z^f}\ U(z^f) - p^{wT} z^f \tag{3.5}$$

where we write ^{T} for the vector transpose. Denoting first and second derivatives of U by U' and U'', respectively, this leads, since z^f unconstrained in (3.5), to the simple equality:

$$p^w = U'(z^f) \tag{3.6}$$

meaning that China's direct impact on world prices is reflected:

$$\frac{\partial p^w}{\partial z^f} = U''(z^f) \tag{3.7}$$

which by concavity is a negative definite symmetric matrix. Hence a higher curvature amounts to a stronger impact. Of course, this is only the direct impact of a change in z^f. The full impact of a change in conditions as represented by a policy scenario will take this function as given and account also for China's reaction to a change in world price.

Furthermore, linearity with respect to the numeraire good implies that the optimization becomes indifferent as to the distribution across consumers and the foreign sector of the demand for the numeraire good. Therefore, noting that surplus maximization leads to $x_{iK} = 0$, because x_{iK} does not contribute to it, we drop explicit representation of numeraire consumption x_{iK} and subsume it under w jointly with public demand for consumption and investment purposes. This leads to the welfare criterion:

$$W = \sum_r \left(\sum_{i \in I_r^u} \upsilon_i(x_i) + \sum_{i \in I_r^r} \upsilon_i(x_i) \right) + \bar{p}_K w + U(z^f) \tag{3.8}$$

(3) The general equilibrium model comprises a detailed component for agricultural production. This component is developed further in section 4 below. Here we merely represent it by defining q_c as the output and e_c the input vector in the transformation function $F_c(q_c, e_c)$, which is strictly quasiconvex non-decreasing in $(q_c, -e_c)$, and appears as constraint

$$F_c(q_c, e_c) \le 0 \tag{3.9}$$

that is the only restriction of the profit maximization at county level, which as will be seen agrees with separate profit maximization for every farm type in every county.

(4) Given net endowments ω_r are introduced at regional level to represent net supply by the non-farm sector supply exogenously, for the reasons mentioned in section 3.1 above.

(5) Every region *r* has commodity balance:

$$\sum_{i \in I_r} x_i + \sum_{r'} v_{rr'} + \sum_{c \in C_r} z_c^+ + g_r \delta^K + m_r^- = \sum_{r'} v_{r'r} + \sum_{c \in C_r} z_c^- + m_r^+ + \omega_r \tag{3.10}$$

where

g_r	scalar denoting use of numeraire good for consumption and investment and as input in trade and transportation
δ^K	*K*-dimensional vector with unit value in the *K*-th row and zero elsewhere

while all other variables are non-negative vectors of dimension *K*:

m_r^+ , m_r^- imports from abroad into region r, and corresponding exports
$v_{rr'}$ flow from region r to r'
x_i consumption of group i
z_c^+ , z_c^- gross commodity inflow into rural areas of county c, and corresponding outflow
ω_r exogenously given net supply by non-agricultural sector

(6) At world level a commodity balance has to hold as well:

$$\sum_r (m_r^+ - m_r^-) + z^f \leq 0 \tag{3.11}$$

(7) To the regional balance a county balance is linked that refers to the rural areas only:

$$e_c + z_c^- = q_c + z_c^+ \tag{3.12}$$

where K-dimensional vectors e_c and q_c refer to the input and output of the transformation function $F_c(q_c , e_c)$ of the rural area in county c.

To inflows z_c^+ and outflows z_c^- are associated the unit transport costs τ_c^+ and τ_c^-, respectively. Similarly, ζ_r^+ and ζ_r^- denote the transport and processing requirements for international trade from the border to the region and vice-versa. Costs of transport between regions are represented through the K-dimensional vectors $\theta_{rr'}$ of inter-regional transport requirements, i.e. demand for non-agriculture as input. We also assume that numeraire demand w in (3.8) is distributed in fixed shares η_r across regions, in every year of simulation. This is for accounting purposes only, as this distribution does not impact on other outcomes, because we disregard trade and transportation margins on the numeraire good. Consequently, total demand g_r of numeraire good K for use in inter- and intraregional transport, as well as for consumption and investment demand (private and public) is:

$$g_r = \sum_{r'} \theta_{rr'}^T v_{rr'} + \sum_{c \in C_r} (\tau_c^{+T} z_c^+ + \tau_c^{-T} z_c^-) + \zeta_r^{+T} m_r^+ + \zeta_r^{-T} m_r^- + \eta_r w \tag{3.13}$$

Basic version

Combining these elements, we obtain as basic version of the welfare model:

$$V^* = \max_{v_{rr'} \geq 0; e_c, m_r^-, m_r^+, q_c, x_i, z_c^-, z_c^+ \geq 0, g_r, w, z^f} \sum_i \upsilon_i(x_i) + \overline{p}_K w + U(z^f)$$

subject to

$$\begin{aligned}
&\textstyle\sum_{i\in I_r} x_i + \sum_{r'} v_{rr'} + \sum_{c\in C_r} z_c^+ + g_r \delta^K + m_r^- = \sum_{r'} v_{r'r} + \sum_{c\in C_r} z_c^- + m_r^+ + \omega_r && (p_r)\\
&g_r = \textstyle\sum_{r'} \theta_{rr'}^T v_{rr'} + \sum_{c\in C_r} (\tau_c^{+T} z_c^+ + \tau_c^{-T} z_c^-) + \zeta_r^{+T} m_r^+ + \zeta_r^{-T} m_r^- + \eta_r w && \\
&\textstyle\sum_r (m_r^+ - m_r^-) + z^f \leq 0 && (p^w)\\
&e_c + z_c^- = q_c + z_c^+ && (p_c)\\
&F_c(q_c, e_c) \leq 0 &&
\end{aligned} \tag{3.14}$$

where we note that the model only determines the total $w + z_K^f$, while split between z_K^f and w is determined afterwards so as to meet a pre-specified net trade surplus for China.

The formulation illustrates the major practical advantage of the welfare approach that it can accommodate a complex economic system in a transparent way, for example, through its first-order conditions, under the prevailing non-negativity bounds on the prices and quantities, the relationships:

$$p_{r'} \leq p_r + p_{rK}\theta_{rr'} \qquad \textit{with equality if } v_{rr'} > 0 \tag{3.15a}$$

$$p_r \leq p^w + p_{rK}\zeta_r^+ \qquad \textit{with equality if } m_r^+ > 0 \tag{3.15b}$$

$$p_r \geq p^w - p_{rK}\zeta_r^- \qquad \textit{with equality if } m_r^- > 0 \tag{3.15c}$$

$$p_c \leq p_r + p_{rK}\tau_c^+ \qquad \textit{with equality if } z_c^+ > 0 \tag{3.15d}$$

$$p_c \geq p_r - p_{rK}\tau_c^- \qquad \textit{with equality if } z_c^- > 0\,. \tag{3.15e}$$

3.3 Specification of the welfare model: full version

The full version of the static welfare model requires some further extensions with additional features that we introduce one-by-one.

Further extensions

(8) We explicitly introduce given population numbers n_i for China. Consumption x_i will now denote the associated per capita consumption vectors. This changes the welfare objective to:

$$W = \sum_i n_i \upsilon_i(x_i) + \overline{p}_K w + U(z^f) \tag{3.16}$$

(9) Next, we consider quantity constraints and price wedges imposed via policies. At regional level we include tariffs ξ_r^+, ξ_r^- and quotas $\underline{m}_r^+, \overline{m}_r^+, \underline{m}_r^-, \overline{m}_r^-$ on international trade, and consumer taxes ξ_i^x, for agricultural products only. Excise tax per quantity unit ξ_c^q is imposed on county production. Lower bounds on exports and import depict state trade. As shown in Ginsburgh and Keyzer (2002, chapter 5), quotas are readily incorporated as bounds on flows

$\underline{m}_r^+ \leq m_r^+ \leq \overline{m}_r^+$, $\underline{m}_r^- \leq m_r^- \leq \overline{m}_r^-$, while tariffs and other taxes on trade can be represented via additional terms in the objective. The resulting "distorted welfare" criterion $\tilde{W}$ is now:

$$\tilde{W} = \sum_i n_i \upsilon_i (x_i) + \overline{p}_K w + U(z_f) - \sum_r (\xi_r^{+T} m_r^+ + \xi_r^{-T} m_r^-) - \sum_c \xi_c^{qT} q_c - \sum_i n_i \xi_i^{xT} x_i \quad (3.17)$$

(10) The next extension is to account for the trade and transportation margins that rural consumers have to face due to their longer distance from the market. Since net buying counties face higher prices because of this and net selling counties lower prices and we only distinguish consumers at regional level, we suppose that every region has a net buying and a net selling part, each with a different regional price, and assume that a fixed fraction Γ_{rk} of total consumption is from net buying rural counties ($c \in C_{rk}^{\circ}$) and hence the remaining fraction $1 - \Gamma_{rk}$ from net selling counties ($c \in C_{rk}^{\bullet}$) for commodity k in region r. These fractions are used to determine the processing requirements, and associated to these the price margins. This is a simplified representation in that the buying-selling classification is only changed via the scenario.

In fact, trade and transportation costs between regional level and rural areas now consist of two components. The first component operates via county-specific factors $\breve{\kappa}_{ck}$ (on farm output) and $\hat{\kappa}_{ck}$ (on farm input), and the second one via regional coefficients τ_r^+ and τ_r^- for, respectively, net buying counties (price increasing) and net selling counties (price decreasing). To implement the latter we define z_{rk}^+ and z_{rk}^- as the net commodity inflows into buying rural counties and the net commodity outflows from selling rural counties:

$$z_r^+ = \Gamma_r \sum_{i \in I_r^r} x_i n_i + e_r^{\circ} - q_r^{\circ} \quad (3.18a)$$

and

$$z_r^- = q_r^{\bullet} - e_r^{\bullet} - (I - \Gamma_r) \sum_{i \in I_r^r} x_i n_i \quad (3.18b)$$

Here, Γ_r is the $K \times K$ diagonal matrix with elements Γ_{rk}, whereas vectors q_r° and $q_r^{\bullet}$ denote gross output of net buying respectively net selling counties, and vectors e_r° and $e_r^{\bullet}$ gross input of net buying respectively net selling counties, in region r. Please note that these definitions of z_{rk}^+ and z_{rk}^- are more specific than the earlier sums $\sum_{c \in C_r} z_{ck}^+$ and $\sum_{c \in C_r} z_{ck}^-$ in (3.14) in that they take into account also rural consumption.

With these definitions the regional commodity balances become:

$$\sum_{i \in I_r^u} x_i n_i + \sum_{r'} v_{rr'} + g_r \delta^K + m_r^- + z_r^+ = \sum_{r'} v_{r'r} + m_r^+ + \omega_r + z_r^- \quad (3.19)$$

Thus, the gap between regional demand (urban consumption, public demand and trade and transportation requirements) and net regional supply is met by inflow from and outflow to rural areas, other regions and abroad. The trade and transportation demand between regional market and counties that is in the basic version equal to $\sum_{c \in C_r} (\tau_c^{+T} z_c^+ + \tau_c^{-T} z_c^-)$, changes into:

$$\tau_r^{+T} z_r^+ + \tau_r^{-T} z_r^- + \sum_{c \in C_r} (\breve{\kappa}_c^{\,T} q_c + \widehat{\kappa}_c^{\,T} e_c) \tag{3.20}$$

Below, after the specification of the full model version, the consequences of these changes for the relations between county prices, market prices and foreign trade prices are derived. In this respect, we emphasize that rural consumption, although consisting of a net buying and a net selling part, is not specified itself at county level. Therefore, there are no county commodity balances in this set-up, which implies that county prices for outputs and inputs may differ. They are denoted by $\breve{p}_c$ and $\widehat{p}_c$, respectively. These prices enter the model of the representative producers in each county *c*, who solve:

$$\begin{aligned} &\Pi_c(\breve{p}_c , \widehat{p}_c) = \max_{e_c , q_c \geq 0} \breve{p}_c{}' q_c - \widehat{p}_c{}' e_c \\ &\qquad \text{subject to} \\ &\qquad\quad F_c(q_c , e_c) \leq 0. \end{aligned} \tag{3.21}$$

and collectively form the supply component.

Full version

The basic version (3.14) of the welfare program now extends to full version:

$$\begin{aligned} V^* = \ &\max_{v_{rr'} \geq 0; e_c, e_r^{\circ}, e_r^{\bullet}, m_r^-, m_r^+, q_c, q_r^{\circ}, q_r^{\bullet}, x_i \geq 0; z_r^-, z_r^+ \geq 0, g_r, w, z^f} \sum_i n_i \upsilon_i (x_i) + \bar{p}_K w + U(z^f) \\ &\qquad - \sum_r (\xi_r^{+T} m_r^+ + \xi_r^{-T} m_r^-) - \sum_i n_i \xi_i^{xT} x_i - \sum_c \xi_c^{qT} q_c \end{aligned}$$

subject to (3.22)

$$\begin{aligned} &\sum_{i \in I_r^u} x_i n_i + \sum_{r'} v_{rr'} + g_r \delta^K + m_r^- + z_r^+ = \sum_{r'} v_{r'r} + m_r^+ + \omega_r + z_r^- && (p_r) \\ &g_r = \sum_{r'} \theta_{rr'}{}^T v_{rr'} + \tau_r^{+T} z_r^+ + \tau_r^{-T} z_r^- + \zeta_r^{+T} m_r^+ + \zeta_r^{-T} m_r^- + \sum_{c \in C_r} (\breve{\kappa}_c^{\,T} q_c + \widehat{\kappa}_c^{\,T} e_c) + \eta_r w \\ &\sum_r (m_r^+ - m_r^-) + z^f \leq 0 && (p^w) \\ &\underline{m}_r^+ \leq m_r^+ \leq \overline{m}_r^+ \\ &\underline{m}_r^- \leq m_r^- \leq \overline{m}_r^- \\ &z_r^+ = \Gamma_r \sum_{i \in I_r^r} x_i n_i + e_r^{\circ} - q_r^{\circ} && (-\psi_r^{\circ}) \\ &z_r^- = q_r^{\bullet} - e_r^{\bullet} - (I - \Gamma_r) \sum_{i \in I_r^r} x_i n_i && (\psi_r^{\bullet}) \\ &e_{rk}^{\circ} = \sum_{c \in C_{rk}^{\circ}} e_{ck} \\ &q_{rk}^{\circ} = \sum_{c \in C_{rk}^{\circ}} q_{ck} \\ &e_{rk}^{\bullet} = \sum_{c \in C_{rk}^{\bullet}} e_{ck} \\ &q_{rk}^{\bullet} = \sum_{c \in C_{rk}^{\bullet}} q_{ck} \\ &F_c(q_c , e_c) \leq 0 \end{aligned}$$

This is the full specification of the Chinagro welfare model with the qualification that the transformation functions F_c will be made explicit in the next section. Welfare equilibrium of model (3.22) generates bounded prices p^w, p_r and margins $\psi_r^{\circ}, \psi_r^{\bullet}$ that support an equilibrium in which consumers and the rest of the world maximize their surplus according to (3.4) and (3.5), respectively, while producers maximize their profits according to (3.21). The equilibrium is unique, and once all distortions are eliminated yields a Pareto efficient solution.

Assuming that foreign trade bounds not binding and that flows z_{rk}^{+} and z_{rk}^{-} are positive, the price relations (3.15) of the basic version are, again under prevailing non-negativity bounds on variables, replaced by:

$$p_{r'} \le p_r + p_{rK}\theta_{rr'} \qquad \textit{with equality if } v_{rr'} > 0 \qquad (3.23a)$$

$$p_r \le p^w + p_{rK}\zeta_r^{+} + \xi_r^{+} \qquad \textit{with equality if } m_r^{+} > 0 \qquad (3.23b)$$

$$p_r \ge p^w - p_{rK}\zeta_r^{-} - \xi_r^{-} \qquad \textit{with equality if } m_r^{-} > 0 \qquad (3.23c)$$

$$\breve{p}_c = p_r + p_{rK}\tau_r^{+} - p_{rK}\breve{\kappa}_c - \xi_c^{q} \qquad \textit{for } c \in C_r^{\circ} \qquad (3.23d)$$

$$\breve{p}_c = p_r - p_{rK}\tau_r^{-} - p_{rK}\breve{\kappa}_c - \xi_c^{q} \qquad \textit{for } c \in C_r^{\bullet} \qquad (3.23e)$$

$$\widehat{p}_c = p_r + p_{rK}\tau_r^{+} + p_{rK}\widehat{\kappa}_c \qquad \textit{for } c \in C_r^{\circ} \qquad (3.23f)$$

$$\widehat{p}_c = p_r - p_{rK}\tau_r^{-} + p_{rK}\widehat{\kappa}_c \qquad \textit{for } c \in C_r^{\bullet} \qquad (3.23g)$$

Hence, the county level prices for both farm outputs ($\breve{p}_{ck}$) and farm inputs ($\widehat{p}_{ck}$) are obtained from regional prices and price margins, depending on whether the county is a net buyer ($c \in C_{rk}^{\circ}$) or a net seller ($c \in C_{rk}^{\bullet}$) of the product concerned. Since we only measure either $\widehat{\kappa}_c$ or $\breve{\kappa}_c$ we make the assumption that $\widehat{\kappa}_c = -\breve{\kappa}_c$. This means that, apart from the producer tax, the same county prices are applied to production and input demand. The margins may be negative as they actually reflect differences in composition, i.e. differences in quality and making of the commodity as well as differences in location inside the region.

Furthermore, the consumer prices of the agricultural commodities can be expressed as marginal utilities that obey for non-negative x_i:

$$\partial\upsilon_i(x_i)/\partial x_i \le p_r + \xi_i^{x} \qquad \textit{for } i \in I_r^{u}\textit{, with equality if } x_i > 0 \qquad (3.24a)$$

$$\partial\upsilon_i(x_i)/\partial x_i \le p_r + \xi_i^{x} + p_{rK}(\Gamma_r\tau_r^{+} - (1-\Gamma_r)\tau_r^{-}) \quad \textit{for } i \in I_r^{r}\textit{, with equality if } x_i > 0 \qquad (3.24b)$$

In fact, we assume that $\Gamma_{rk}\tau_{rk}^{+} = (1-\Gamma_{rk})\tau_{rk}^{-}$, in order to maintain the interpretation of p_r (which is actually observed) as mean consumer price for rural areas. One may also note from (3.24) that consumer prices are equal across all consumer groups, apart from the consumer tax. This specification is possible since quality differences are treated as additional consumption of commodity *K* (non-agriculture).

Regarding functional forms domestic consumers are taken to obey a Linear Expenditure System, with a utility function:

$$u_i(x_i) = \gamma_{i0} \prod_{k=1}^{K-1} (x_{ik} - \underline{x}_{ik})^{\gamma_{ik}} + \overline{p}_K x_{iK} \tag{3.25}$$

where $\sum_{k=1}^{K-1} \gamma_{ik} < 1,\ \gamma_{ik} \geq 0$, while foreign demand has utility

$$U(z^f) = \sum_{k=1}^{K-1} \beta_k^f z_k^f - \sum_{k=1}^{K-1} \gamma_k^f (z_k^f)^2 + \overline{p}_K z_K^f \tag{3.26}$$

which leads to a linear price function (3.6). International price trends beyond the influence of China can be represented in this function via exogenous parameter shifts.

The model is solved by decomposition into a supply component that takes all prices as given and generates net supply from (3.21), a world price component that generates international prices on the basis of (3.6), and an exchange component that takes net supply and world prices as given and generates prices as Lagrange multipliers. Prices are adjusted iteratively on that basis. The model owes its computational strength to particular features of this decomposition addressed briefly in section 5.

4. The agricultural supply component

4.1 Embedding in the welfare program

The county-level specification of the farm sector is embedded within the welfare program of the previous section via the trade links of farm types to regional markets. To make this explicit, we return to transformation function constraint (3.9), disaggregating it further by distinguishing different land use types. For every land use type $j = 1,...,J$ one representative farm is considered in each county, with a transformation function of its own.

Every county-level transformation function F_c can now be interpreted as the value function in the usual transformation function construct:

$$\begin{aligned} & F_c(q_c, e_c) = min_{e_{cj}, q_{cj} \geq 0, F} \; F \\ & \qquad subject\ to \\ & \qquad \quad F \geq F_{cj}(q_{cj}, e_{cj}) \\ & \qquad \quad \sum_j e_{cj} \leq e_c \\ & \qquad \quad \sum_j q_{cj} \geq q_c \end{aligned} \tag{4.1}$$

where q_{cj} and e_{cj} denote the output and input vector of land use type j in county c, respectively, and all functions F_{cj} are strictly quasi-convex non-decreasing in $(q_{cj}, -e_{cj})$. For notational ease, we did refer to value function F_c in the welfare program itself, but it is clear from (4.1) that this is equivalent to introducing the constraints:

$$\begin{aligned} & e_{cj}, q_{cj} \geq 0 \\ & F_{cj}(q_{cj}, e_{cj}) \leq 0 \\ & \sum_j e_{cj} \leq e_c \\ & \sum_j q_{cj} \geq q_c \end{aligned} \tag{4.2}$$

In the previous section we defined already the county-specific prices for farm outputs ($\breve{p}_{ck}$) and farm inputs ($\hat{p}_{ck}$), and showed their relations to the regional prices in equations (3.23d) – (3.23g). As mentioned, these prices enter model (3.21) of the representative producers in each county *c*, which becomes with explicit mentioning of the land use types:

$$\begin{aligned} & \Pi_c(\breve{p}_c, \hat{p}_c) = \; max_{e_{cj}, q_{cj} \geq 0} \sum_j \left(\breve{p}_c^{\,T} q_{cj} - \hat{p}_c^{\,T} e_{cj} \right) \\ & \qquad subject\ \ to \\ & \qquad \quad F_{cj}(q_{cj}, e_{cj}) \leq 0. \end{aligned} \tag{4.3}$$

The present section presents the details of this county supply program. As is well known, if all farmers take prices as given, profit maximization at sector (here county) level is equivalent to independent profit maximization by all farms separately.

4.2 General structure

We now consider farms as representative producers indexed j, explicitly add labor and, dropping the county index for ease of notation, change program (4.3) to:

$$
\begin{aligned}
\Pi(\breve{p},\hat{p}) = \ & max_{e_{kj},q_{kj}\geq 0,L_j\geq 0} \sum_j (\sum_k \breve{p}_k q_{kj} - \sum_k \hat{p}_k e_{kj}) \\
& subject\ \ to \\
& \quad F_j(q_{1j},...,q_{Kj},e_{1j},...,e_{Kj},L_j,\bar{A}_j) \leq 0 \\
& \quad \sum_j L_j \leq \bar{L} \qquad\qquad\qquad\qquad (\mu)
\end{aligned} \tag{4.4}
$$

where q_{kj} stands for the output of k from land use j and e_{kj} for input, where L_j is the labor applied to land use type j and $\bar{A}_j$ the given number of capacity units (area or number of stable units, depending on the land use type). The shadow price of labor is indicated between brackets, as μ.

In the individual maximization programs of the equivalent formulation referred to above, labor costs μL_j are also subtracted from the objective, leaving only the transformation function and non-negativity conditions as constraints whereas the county labor balance $\sum_j L_j = \bar{L}$ connects the individual programs via adjustment of shadow price μ.

Let the set of representative farms be subdivided in three classes with index sets J_p, J_g and J_o, representing a price responsive class, grazing, and other land use activities respectively:

(i) *Price responsive class,* $j \in J_p$: yield is dependent on labor and equipment intensity per hectare for crops and per stable unit for livestock, and labor is allocated flexibly between different uses:

Irrigated cropping
Rainfed cropping
Specialized dairy farm
Traditionally mixed nonruminant farm
Intensified nonruminant farm

(ii) *Grazing,* $j \in J_g$: for grazing, yield is price independent and depends on the given livestock density (herdsize per hectare) and on the supplementary feed provided per hectare. The sector's output is price responsive, nonetheless, in that we allow for substitution between the commodities entering the aggregate yield index. Labor use per unit $\bar{A}_j$ is fixed.

(iii) *Other land use types,* $j \in J_o$: these have fixed input ε_{kj} and output o_{kj} of various commodities k per unit $\bar{A}_j$; labor use per unit $\bar{A}_j$ is fixed as well:

Tree cropping
Draught animal system
Traditionally mixed ruminant farm

Also five farm-related sectors (machine power, household waste, household manure, green feed and utilizable grass) are treated as part of J_o.

On the output side, the land use types are further subdivided into activities *h*. For cropping these activities are defined in terms of crops that are subsequently processed into main products and byproducts, and for livestock in terms of animal types, also processed into main and by-products. The list of activities by land use type is given in Appendix A.

In addition to the tradable commodities indexed *k*, the land use types supply or use local commodities only traded inside the own county, such as manure, crop residuals and animal power. Use of these local commodities follows fixed coefficients, bilaterally specified per unit of $\overline{A}_j$ in which *j* may refer either to the delivering or the receiving land use type (see section 4.7).

The exogenous capacities $\overline{A}_j$ play an important role in the model. In simulation, these capacities follow directly from scenario assumptions on resources that follow a slightly different classification, listed in Appendix A, which, however, is not essential for an understanding of the specification and operation of the supply module and is, therefore, not considered in the present section.

4.3 Separability

Revisiting the steps in Albersen et al. (2002), we further structure the transformation functions F_j. First, we impose separability between input and outputs:

$$F_j(q_{1j},...,q_{Kj},e_{1j},...,e_{Kj},L_j,\overline{A}_j) = H_j(q_{1j},...,q_{Kj}) - G_j(e_{1j},...,e_{Kj},L_j,\overline{A}_j) \qquad (4.5)$$

This separability assumption has the disadvantage that it takes inputs to impact only on aggregate output G_j of the land use type. We introduce it for a very practical reason. While it is empirically already very demanding to develop a data set with commodity inputs by land use type, further split to different outputs within the same land use type is seldom possible. Be this as it may, the separabality assumption also keeps the model structure more transparent and more easily amenable to decomposition.

Regarding the properties of the functions in (4.5) we impose for the price responsive class as well as for grazing:

Assumption TO (transformation function, outputs): Output function H_j is (i) strictly quasiconvex, convex; (ii) nondecreasing; and (iii) homogeneous of degree one.

Assumption TI (transformation function, inputs): Input function G_j is (i) concave; (ii) nondecreasing; and (iii) homogeneous of degree one, and (iv) strictly concave increasing in $(e_{1j},...,e_{Kj},L_j)$.

Homogeneity assumption TO(iii) makes it possible to specify the output of separate crops relative to an (independently defined) aggregate output index Y_j. Function G_j will be measured in the same aggregate unit. Homogeneity assumption TI(iii) makes the optimal commodity and labor input allocation independent of the level of resource $\bar{A}_j$. Therefore, the optimal allocation can be derived per unit of $\bar{A}_j$. In fact, the function G_j will be specified as a function with two branches, one for labor (equipped with machinery) and one for commodity inputs, which are taken to be expressible via an overall input index.

The separability assumption makes it possible to decompose program (4.4) into the following three components:

(1) Optimal output composition for each land use type: section 4.4.
(2) Optimal inputs and optimal aggregate output for each land use type, at given shadow prices for labor μ: section 4.5.
(3) Shadow prices for labor μ clearing the balance of agricultural labor: section 4.6.

4.4 Optimal output composition and the revenue index

For the price responsive class and for grazing, homogeneity property TO(iii) enables us to write the aggregate output value as the product of the aggregate quantity index Y_j and aggregate unit revenue index R_j. Commodity output q_{kj} of land use type j is without loss of generality expressed as the product:

$$q_{kj} = g_{kj} Y_j \tag{4.6}$$

Hence, the variable g_{kj} is the volume of net output of commodity k per unit of aggregate output Y_j. Revenue maximization seeks an optimal composition of this aggregate,

$$\begin{aligned} R_j &= max_{g_{kj} \geq 0} \textstyle\sum_k \breve{p}_k g_{kj} \\ &\quad subject\ \ to \\ &\qquad H_j(g_{1j},...,g_{Kj}) \leq 1 \end{aligned} \tag{4.7}$$

where we define H_j as $H_j(g_{1j},...,g_{Kj}) = max(1 + h_j(g_{1j},...,g_{K-1,j}) + g_{Kj} , \sum_{k=1}^{K-1} \vartheta_{kj} g_{kj})$ for convex, strictly quasiconvex, increasing , homogeneous of degree one function h_j and fixed non-negative coefficients ϑ_{kj}. The second branch of this definition could be any convex function that is homogeneous of degree one. We isolate g_{Kj} because it refers to the input needed to manage cropping patterns and rotation; it is non-positive.

As mentioned in section 4.2, the actual implementation does not directly formulate the optimal allocation in terms of commodities k. Rather it distinguishes for each land use type several activities h (crops, animals) that compete for resource $\bar{A}_j$. The activities have yields y_{hj} per unit of area (or stable capacity) that are linked to the development of aggregate yield $y_j = Y_j / \bar{A}_j$ of the land use type, whereas fixed output coefficients B^j_{kh} represent the output of commodities *k* (e.g. rice, milk, meat) per unit of output of activity *h* (e.g. paddy, milk cows, pigs) in land use type *j*.

Thus, unit revenue R_j is obtained from optimal land (or stable capacity) allocation within every land use type, and for this we specify the program, for real cost function C_j replacing h_j in (4.7) and with ϑ-coefficients at unity:

$$R_j(\rho_j, \hat{p}_K) = max_{\alpha_{hj} \geq 0} \left(\sum_h \rho_{hj}\alpha_{hj} - \hat{p}_K C_j(\alpha_{1j},...,\alpha_{Hj}) \right) / y_{0j}$$
$$subject\ to \qquad (4.8)$$
$$\sum_h \alpha_{hj} \leq 1$$

where α_{hj} the area share that measures the fraction of $\overline{A}_j$ assigned to activity *h*, and revenue $\rho_{hj}(\breve{p}) = r_{hj}(\breve{p})\overline{y}_{hj}$, with $r_{hj}(\breve{p}) = \sum_k max(B^j_{kh}\breve{p}_k, 0)$ for activity *h* in land use type *j*. Function C_j represents the rotation costs on land use type *j*. Like function h_j, it has to be homogeneous of degree one, strictly quasi-convex, convex and increasing. The homogeneity property is the feature that makes it possible to formulate the problem in terms of area shares. Strict quasi-convexity rules out linearity and, thereby, full specialization. Note that the division by reference aggregate yield y_{0j} ensures that optimal revenue $R_j(\rho_j, \breve{p}_K)$ is expressed as revenue per unit of aggregate output.

The output price $\breve{p}_k$ at county level is related to the market prices p_r of the corresponding region, according to price relationships (3.23d) – (3.23g) from region to county. Recall that $\breve{p}_k$ could be negative because of processing cost but this is not an issue here, as we may assume free disposal of any negatively priced output *k* from activity *h*, and there always will be some positively priced activity.

Per hectare (or per unit of stable capacity) yield y_{hj} is based on exogenous levels $\overline{y}_{hj}$ that adapt by a factor common across all crops of farm type *j* to changes in aggregate yield y_j:

$$y_{hj} = \overline{y}_{hj} \frac{y_j}{y_{0j}}. \qquad (4.9)$$

Hence, crop-specific adaptations in yield can only be accommodated via exogenous shifts. This assumption has been implemented above in the specification of the objective of program (4.8) that, otherwise, would depend on the ratio of endogenous variables y_{hj} / y_j. The yield adaptations can now be attributed either to changes in cropping intensity or to changes in the yields of a single harvest, as shown explicitly in the tabulation of the model outcomes.

Rotation cost function C_j obeys

$$C_j(\alpha_{1j},...,\alpha_{Hj}) = -\sum_h \tau_{hj}\alpha_{hj} + D_j(\alpha_{1j},...,\alpha_{Hj}) \qquad (4.10)$$

where τ_{hj} are the direct implicit returns or costs (if τ_{hj} is negative) for crop *h*, and D_j is a standardized function, proportionate by a factor κ_j to the ℓ_2-norm:

$$D_j(\alpha_{1j},...,\alpha_{Hj}) = \kappa_j(\sum_h(\alpha_{hj})^2)^{\frac{1}{2}} \quad (4.11)$$

Turning to the solution, it follows from the envelope theorem that optimal land allocation satisfies:

$$\alpha_{hj} = \frac{\partial R_j(\rho_j, \hat{p}_K)}{\partial \rho_{hj}}. \quad (4.12)$$

Unlike in usual CES-based formulation, this specification of the rotation cost function allows for smooth but complete switching in and out of activities. Hence, depending on the relative profitability some activities may appear while others vanish. From this solution, commodity outputs g_{kj} per unit of aggregate output are derived as:

$$g_{kj} = \frac{1}{Y_j}\sum_h B^j_{kh} y_{hj}\alpha_{hj}\bar{A}_j - \partial_k \frac{C_j(\alpha_{1j},...,\alpha_{Hj})}{y_{0j}} \quad (4.13)$$

$$with\ \partial_k = 0\ for\ k = 1,...,K-1\ and\ \partial_k = 1\ for\ k = K$$

Hence, for $j \in J_p$ and $j \in J_g$ the relation between aggregate output and commodities consists of two steps: an optimal land or stable capacity allocation, and fixed coefficients from activity levels to commodities. For the other land use types, $j \in J_o$, both steps are governed by fixed coefficients. Please note that the introduction of activities and the B-matrix has changed the definition of g_{kj} for $k = K$ compared to its use in program (4.7).

For later reference we also define composite revenue per unit of aggregate output (scale-independent) directly as function of selling prices $\breve{p}$:[2]

$$r_j(\breve{p}) \equiv R_j(\rho_j(\breve{p}), \hat{p}_K) \quad (4.14)$$

4.5 Optimal input demand and aggregate output

We now turn to the demand for inputs associated to the aggregate output of land use types $j \in J_p$ and $j \in J_g$. Homogeneity property TI(iii) makes the optimal input-output structure independent of the level of resource use $\bar{A}_j$. We denote labor use per unit of $\bar{A}_j$ by ℓ_j. Hence, $\ell_j = L_j / \bar{A}_j$. Earlier, we defined already aggregate yield as $y_j = Y_j / \bar{A}_j$.

[2] Since $\hat{p}_K$ is equal to $\breve{p}_K$ and, anyhow, exogenous, it is omitted as argument in the definition of r_j.

For simplicity of notation we limit attention to a single purchased input, denoted as function $F_j^e(e_{1j},...,e_{Kj})$, to be referred to as feed and purchased at the given price $p_j^f(\hat{p})$ albeit that the formulation applies to an arbitrary number of current inputs with purchasing prices $\hat{p}_k$.

Input function $G_j(e_{1j},...,e_{Kj},L_j,\overline{A}_j)$ becomes a composite function that can be denoted $\tilde{G}_j(F_j^e(e_{1j},...,e_{Kj}),L_j,\overline{A}_j)$. The simplest formulation keeps constant returns in input demand:

$$\begin{aligned} p_j^f(\hat{p}) = min_{e_{kj}\geq 0} \textstyle\sum_k \hat{p}_k e_{kj} \\ subject\ to \\ F_j^e(e_{1j},...,e_{Kj}) \geq 1 \end{aligned} \tag{4.15}$$

where functions F_j^e concave, differentiable and linear homogeneous. Unit cost function p_j^f is a concave, differentiable and linear homogeneous function dependent on input prices $\hat{p}$. We denote by f_j the optimal level of aggregate input per unit of the resource, hence $f_j = F_j^e(e_{1j},...,e_{Kj})/\overline{A}_j$ for vectors e_j with optimal composition. For every commodity k optimal input demand can be obtained via Shephard's Lemma as:

$$e_{kj} = \frac{\partial p_j^f(\hat{p})}{\partial \hat{p}_k} f_j \overline{A}_j. \tag{4.16}$$

It may be restrictive to assume a yield-independent composition of this purchased input but this formulation has the advantage that it maintains continuity of the input demand curve. Nonetheless, extensions can be envisaged in which the input function p_j^f and its derivative to $\hat{p}$ depend on yield.

Using this aggregate input price function $p_j^f(\hat{p})$ and the aggregate revenue function $r_j(\breve{p})$, optimal inputs and aggregate output per unit of resource $\overline{A}_j$ follow for land use type *j* from:

$$max_{y_j,f_j,\ell_j\geq 0}\{ r_j(\breve{p})y_j - p_j^f(\hat{p}) f_j - \mu \ell_j \mid y_j \leq g_j(f_j,\ell_j)\} \tag{4.17}$$

where f_j denotes the input aggregate per unit of the resource, $f_j \equiv F_j^e(e_j)/\overline{A}_j$, and g_j the yield function per unit of resource $\overline{A}_j$, $g_j(f_j,\ell_j) \equiv \tilde{G}_j(F_j(e_j/\overline{A}_j),\ell_j,1)$.

In this definition, function $g_j(f_j,\ell_j)$ is the aggregate yield function of land use type *j*. We opt for a specification with two branches,[3] one for the relation between y_j and f_j and one for the relation between y_j and ℓ_j :

[3] Earlier experience in estimating cross sectional agricultural production functions for China as described in Albersen et al. (2002), contributed to this specification.

$$\begin{aligned}&max_{y_j, f_j, \ell_j \geq 0} \quad r_j(\breve{p}) y_j - p_j^f(\hat{p}) f_j - \mu \ell_j \\ &\text{subject to} \\ &y_j \leq g_j^f(f_j) \\ &y_j \leq g_j^\ell(\ell_j)\end{aligned} \tag{4.18}$$

Hence, the minimum of $g_j^f(f_j)$ and $g_j^\ell(\ell_j)$ creates the actual upper bound on aggregate yield y_j. Alternatively, we may write program (4.18) as a mixed primal-dual problem, with a primal formulation for the labor branch, as above, and a dual formulation for the commodity input branch:

$$max_{y_j, \ell_j \geq 0} \quad \{ r_j(\breve{p}) y_j - c_j(p_j^f(\hat{p}), y_j) - \mu \ell_j \mid y_j \leq g_j^\ell(\ell_j) \}, \tag{4.19}$$

where

$$c_j(p_j^f(\breve{p}), y_j) \quad = min_{f_j \geq 0} \{ p_j^f(\breve{p}) f_j \mid g_j^f(f_j) \geq y_j \}. \tag{4.20}$$

For the labor branch of this profit maximization model of the representative county farms we use a *Mitscherlich-Baule (MB) yield function,* with given asymptote (yield potential) $\bar{y}_j$ and with labor as sole input equipped with animal and machine power. Land use types compete for labor. All other inputs are dealt with as “feed” in the commodity input branch: fertilizer and manure for crops and purchased or locally available animal feed for livestock. We use a Mitscherlich-Baule production function because it possesses, like the logistic curve, an asymptote that can accommodate information on the yield potential of each land-use type in every county. Its formulation is as follows:

$$g_j^\ell(\ell_j) = \bar{y}_j(1 - e^{\alpha_j - \beta_j \ell_j}) \tag{4.21}$$

where we assume that $\beta_j > 0$ as additional output requires additional labor. Other restrictions on the parameters are discussed in Assumption Y below.

Regarding the cost function of the purchased “feed” inputs for given yield $y_j = Y_j / A_j$, in principle the only requirement is that the cost function should be *convex non-decreasing* in y_j. However, to maintain a closed form solution we also take the cost function to be *piecewise linear in* y_j, which allows for an arbitrarily close approximation:

$$c_j(p_j^f, y_j) = min_{f_j \geq 0} \{ p_j^f f_j \mid \ f_j \geq \eta_j^i y_j - \gamma_j^i, i = 1, ..., M \} \tag{4.22}$$

where the coefficients η_j^i and γ_j^i are the fixed coefficients. We only need to assume $\eta_j^1 = 0$, $\gamma_j^1 = 0$ (possibility of inaction) and $\eta_j^i \geq 0$ (no free lunch), as the minimization will on this basis select the appropriate constraints on the boundary that lead to a piecewise linear cost function that is concave non-decreasing in prices, and convex non-decreasing in yield. However, to ease the calibration we actually specify the various segments by assuming that $\gamma_j^i > 0$ and $\eta_j^i > 0$ for $i = 2,...,M$, while both the slopes and the intercepts are taken to rise with *i*:

$$\eta_j^i > \eta_j^{i-1} \text{ and } \frac{\gamma_j^i}{\eta_j^i} > \frac{\gamma_j^{i-1}}{\eta_j^{i-1}}, \quad i = 3,...,M \tag{4.23}$$

The intercept γ_j^i makes it possible to account for fixed or exogenous and free supply of the input as a given quantity (manure and natural soil fertility for crops, roughage and hey for ruminants, and household residuals for pork and poultry) and a switch point will occur once an additional input has to be purchased with rising output. This defines the indicator function:

$$i_j(y_j) = \{ i \in \{1,...,M\} \mid \hat{y}_j^i \le y_j \le \hat{y}_j^{i+1} \} \tag{4.24}$$

where $\hat{y}_j^1$ is the yield obtainable with zero feed and zero labor input beyond which labor will be required, $\hat{y}_j^2$ the yield with positive labor and zero feed beyond which feed will be required and , generally, beyond that $\hat{y}_j^{i+1}$ is the yield at the switch point between the branches i and $i+1$ Consequently, in between switching points the unit profit has a constant derivative w.r.t. y_j :

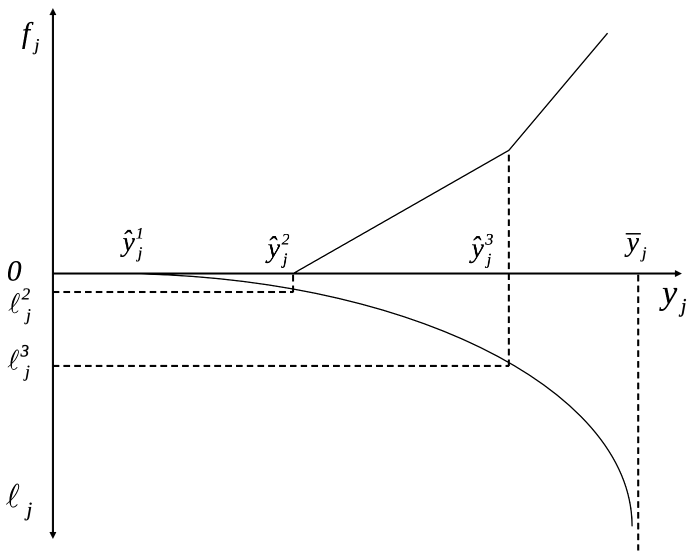
$$\pi_j^i = r_j(\breve{p}) - p_j^f \eta_j^i \tag{4.25}$$

We remark that the zero-labor intercept $\hat{y}_j^1$ corresponds to $y_j(0) = \bar{y}_j(1 - e^{\alpha_j})$. Further assumptions about the relation between the two branches are summarized in Assumption Y:

Assumption Y (first and second input regime): (i) output is positive with zero labor input: $\alpha_j < 0$; (ii) labor has positive marginal productivity: $\beta_j > 0$; (iii) no feed is required at the maximal yield attainable without labor: $\eta_j^1 = 0, \gamma_j^1 = 0, \gamma_j^2 / \eta_j^2 \ge \bar{y}_j(1 - e^{\alpha_j})$.

Under these assumptions both yield branches of program (4.18) are concave functions and have positive yield for zero input. Therefore, the technical relationships are as in Figure 4.1.

Figure 4.1 Two-branch production function of a typical farm *j*



The upper panel of Figure 4.1 indicates how much fertilizer per hectare f_j (respectively, feed per stable unit) is needed to achieve a given yield y_j. The lower panel shows how much yield can be obtained from given labor per hectare (respectively, labor per stable unit). The yield appears on the horizontal yield-axis, the level $\hat{y}_j^1$ refers to the yield obtained without any input of purchased fertilizer and formal labor and bullock or tractor power and equipment (i.e. with only child labor or informal work), $\hat{y}_j^2$ to the yield obtained with formal labor ℓ_j^2 but without purchased fertilizer (resp. purchased feed), i.e. with locally available manure and nightsoil (resp. with locally available crop and household residues), and, finally, $\hat{y}_j^3$ to the yield-threshold beyond which purchased fertilizer (resp. feed) use becomes less effective. The curve in the lower panel shows the response with decreasing returns to labor, where the bio-physical cropping potential $\bar{y}_j$ appears on the right-hand side as a ceiling (asymptote) and measures the maximal amount of biomass that could be produced through photosynthesis under the given climate and soil conditions. For extensive grazing, a similar relationship is postulated but with the herd size in the role of labor input.

4.6 Optimal labor allocation

The final and main step in the decomposition of program (4.4) is to determine the shadow prices μ for agricultural labor by imposing at county-level the labor balance $\sum_j L_j = \bar{L}$. Once these shadow prices are known, the optimal labor input ℓ_j, the optimal "feed " input f_j and the optimal aggregate yield y_j of the land use types follow from program (4.19), the individual commodity inputs e_{kj} from (4.16) and the commodity outputs q_{kj} from (4.6) and (4.13).

Land use types J_p, J_g and J_o obtain labor balance in different ways. For $j \in J_o$ labor demand is exogenous and, therefore, subtracted a priori from the right-hand side of the balance. The same applies to $j \in J_g$, the land use class grazing. Since for this class the "harvesting labor" is essentially performed by the herds themselves, the yield, with associated demand for supplementary feed, follows from the herdsize that is exogenously given. Moreover, the labor of herders cannot be considered substitutable for that of cropping activities because of the large distances involved. Hence, we only consider labor allocation between the price responsive sectors $j \in J_p$, noting that the specification for grazing allows, nonetheless, for price response of the commodity bundle on the output side, following (4.12).

Henceforth, this section only considers the price responsive sectors proper, $j \in J_p$. Labor is a local, non-tradable resource, and $\bar{L}$ is the fixed total labor available for these sectors in a particular county. It is as the input feed f_j per hectare complementary to yield, but explicitly incorporates the yield potential. The profit maximizing labor allocation of a county among land use types $j \in J_p$, corresponds to a sub-problem of program (4.4). It meets a land and labor constraint and is obtained from:

$$max_{A_j, y_j \geq 0}\{ \sum_j \tilde{\pi}_j(\breve{p}, p_j^f, y_j)A_j \mid \sum_j A_j \ell_j(y_j) \leq \bar{L}, A_j \leq \bar{A}_j \} \quad (4.26)$$

where the profit per unit of $\bar{A}_j$ is defined by subtracting the “feed” costs from the revenues:

$$\tilde{\pi}_j(\breve{p}, p_j^f, y_j) = r_j(\breve{p})y_j - c_j(p_j^f, y_j) \quad (4.27)$$

and where the labor requirement function $\ell_j(y_j)$ is the inverse of the *Mitscherlich-Baule* yield function $g_j^{\ell}(\ell_j)$, discussed in the previous section. This function is increasing and therefore has a continuous inverse, in closed form. Clearly, $\ell_j(y_j)$ is convex increasing. If for some land use type, net profit $\tilde{\pi}_j$ remains positive for any y_j, all labor will be employed, and the marginal productivity or wage rate μ will be positive in the optimum.

From Assumption Y follows that $A_j = \bar{A}_j$ is optimal, since production is possible without labor and feed, and the unit revenue $r_j(\breve{p})$ is non-negative by construction. Hence, program (4.26) reduces to the convex program:[4]

$$\begin{aligned} & max_{y_j, f_j, \ell_j \geq 0} \sum_j (r_j(\breve{p})y_j - p_j^f f_j)\bar{A}_j \\ & \textit{subject to} \\ & \qquad y_j \bar{A}_j \leq \bar{y}_j (1 - e^{\alpha_j - \beta_j \ell_j}) \bar{A}_j \\ & \qquad f_j \bar{A}_j \geq (\eta_j^i y_j - \gamma_j^i)\bar{A}_j,\ i = 1,...,M \qquad (\theta_j^i) \\ & \qquad \sum_j \ell_j \bar{A}_j \leq \bar{L} \qquad (\mu) \end{aligned} \quad (4.28)$$

[4] This program can also be derived by inserting the yield functions g_j^{ℓ} and g_j^f for each land use type in program (4.18), scaling it up to $\bar{A}_j$ and including it in county program (4.4).

For positive μ, as labor will not be wasted the MB-restriction holds with equality. For zero μ, yields will for all *j* be at some threshold point $\hat{y}^i_j$, beyond which the marginal cost exceeds marginal revenue and not all available labor will be employed but the MB-restriction can be taken to hold with equality, nonetheless (if marginal cost happens to be equal to marginal revenue, inequality could be optimal as well). Hence, we have the first-order conditions:

(a) $p^f_j \geq \sum_i \theta^i_j \quad \perp \quad f_j \geq 0$

(b) $f_j \geq (\eta^i_j y_j - \gamma^i_j) \quad \perp \quad \vartheta^j_i \geq 0$

(c) $\sum_j \bar{A}_j \ell_j \leq \bar{L} \quad \perp \quad \mu \geq 0$

(d) $\beta_j(r_j(\breve{p}) - \sum_i \theta^i_j \eta^i_j)\bar{y}_j e^{\alpha_j - \beta_j \ell_j} \leq \mu \quad \perp \quad \ell_j \geq 0$

with $y_j = \bar{y}_j(1 - e^{\alpha_j - \beta_j \ell_j})$.

Together conditions (a) and (b) determine the branch selected. In between switch points, we have, by (a) the equality $\theta^i_j = p^f_j$. By contrast, at a switch point between two branches, positive θ^i_j for two *i*-values can co-exist that sum to p^f_j. At such a point these values adjust on an interval of wage rates, so as to keep labor intensity fixed. Condition (d) shows that once the wage is given and the branch of the input function is known, the optimal labor demand follows for each land-use type separately. Moreover, every land use type has a wage-level of its own beyond which no labor is used and $\hat{y}^1_j$ prevails.

We do not detail the algorithm for solving this program and limit ourselves to observing that Assumption Y guarantees in (4.26) continuity of labor allocation functions $\ell_j(\breve{p}, p^f, \bar{L}, \bar{A})$, yield functions $y_j(\breve{p}, p^f, \bar{L}, \bar{A})$ and input demand functions $f_j(\breve{p}, p^f, \bar{L}, \bar{A})$, whereas several components of the decomposition of the previous sections can be used for analytical calculations. In fact, after further reduction of (4.28) the algorithm actually operates on :

$$max_{\ell_j \geq 0} \sum_j \left[r_j(\breve{p})\bar{y}_j(1 - e^{\alpha_j - \beta_j \ell_j}) - p^f_j \; max(max_i(\eta^i_j \bar{y}_j(1 - e^{\alpha_j - \beta_j \ell_j}) - \gamma^i_j),0) \right] \bar{A}_j \qquad (4.29)$$

subject to

$$\sum_j \ell_j \bar{A}_j \leq \bar{L} \qquad (\mu)$$

This program can be solved exactly by means of an algorithm with finite termination. We insist on this property because the outcome is to be used as a module within the algorithm solving the general equilibrium welfare program, and therefore, has to be evaluated often and with good precision. At the same time, the number of counties is large. Besides its speed, an exact algorithm has the advantage of high numerical precision.

Finally, we emphasize that labor in these balances is not merely the (weighted) sum of male, female and child labor employed in agriculture but also accounts for the machinery and animal power used. Therefore, we have used the term 'equipped' labor. This definition makes it difficult to indicate the resulting shadow prices as observable market wages. The role of the labor balances is rather 'coordinating power use across land use types' than 'market clearing'. In this respect, we may also mention that the different land use types also interact via non-marketed local deliveries, which are not visible in the equations above but will be explained in more detail in the next subsection.

4.7 Local deliveries of animal feeds and plant nutrients

So far, all deliveries of animal feeds and plant nutrients are taken to take place at prevailing market prices in the county. Hence, crop residuals used as feed are in the output matrix with elements B^j_{kh} expressed as animal feeds through conversion to usable calories, and manure from livestock to fertilizer equivalents, while accounting for various loss factors. Consequently, these are byproducts whose valuation follows the prices of the goods they compete with. Similarly, nightsoil may be valued as a fertilizer. All land use types supply such goods, generally for use by other land use types.

Yet, we must also account for the fact that these byproducts are only to a limited extent tradable among farmers, and hence remain available freely to the farmers as natural resources without alternative use. For crops, the natural fertility of the soil is usually treated in this way. Rather than being rewarded directly, it is a factor that raises the productivity of the land, and is rewarded via the land rent.

Similarly, manure and feeds from household waste that are available locally can be represented through positive shifters σ^f_j on the intercepts γ^i_j of the feed and fertilizer demand curve. Starting from $i = 2$, the same shifters apply to all branches, implying that the piecewise linear curve is shifted but keeps its shape. For simplicity, we assume that these shifters depend on the areas cultivated and livestock numbers only. Hence the intercepts specified so far become variable, for $i = 2,...,M$:

$$\gamma^i_j = \breve{\gamma}^i_j + \sigma^f_j \tag{4.30}$$

where $\breve{\gamma}^i_j$ are the estimated coefficients that are differentiated by region or agro-ecological zone, while the shifters are by county, treated as data in the estimation process and as variables in model simulation, and defined as:

$$\sigma^f_j = (\textstyle\sum_{j'} \bar{\sigma}^f_{j'j} \bar{A}_{j'}) / \bar{A}_j \tag{4.31}$$

where $\bar{\sigma}^f_{j'j}$ is the supply coefficient of local manure or feed from land use j' to land use *j* per unit of the supplying land use j'. Combined, relationships (4.30) and (4.31) also suggest that the σ-coefficients should to the extent possible reflect marginal (as opposed to average) effects of changes in activity levels $\bar{A}_j$ on the supply of labor and feed and fertilizer substitutes, since for these non-priced goods average values do not play any role. Similarly, for animal and mechanical power that to some extent substitute for labor we introduce the shifter:

$$\alpha_j = \breve{\alpha}_j + \sigma^\alpha_j \tag{4.32}$$

with

$$\sigma^\alpha_j = (\textstyle\sum_{j'} \bar{\sigma}^\alpha_{j'j} \bar{A}_{j'}) / \bar{A}_j \tag{4.33}$$

where $\bar{\sigma}^{\alpha}_{j'j}$ is the supply coefficient of local power from land use j' to land use j, expressed per unit of j'.

Local supplies enter the production relationships in different ways. For local power, we assume that it increases the capacity of individual workers, since it consists of animal and mechanical traction. Hence, we adopt a multiplicative formulation to incorporate it within the Mitscherlich-Baule function that modifies to:

$$y_j = \bar{y}_j(1 - exp(\alpha_j - \beta_j \ell_j(\kappa^1_\alpha + \sigma^\alpha_j)^{\kappa^2_\alpha}) \qquad (4.34)$$

where a positive intercept κ^1_α is needed to ensure that labor without any implements remains productive, while the exponent κ^2_α reflects the decreasing returns to traction with $0 < \kappa^2_\alpha < 1$.

For feed-fertilizer the local inputs act as perfect substitutes as losses and alternative uses such as fuel are already accounted for in the calculation of the σ-coefficients, for $i = 2,..,M$:

$$f_j \geq max(-(\breve{\gamma}^i_j + \sigma^f_j) + \eta^i_j y_j, 0) \qquad (4.35)$$

Here the important point to note is that a high availability of local inputs will eliminate all purchases and tend to take the input demand in the first segment, with possibly unused surpluses of these inputs.

The coefficients $\bar{\sigma}^f_{j'j}$ and $\bar{\sigma}^\alpha_{j'j}$ cover four types of local feeds (straw and other crop residuals, grass, green fodder and household waste), two types of local fertilizer (animal manure and nightsoil) and two types of power (animal and mechanized traction). In the relations above, the land use types j refer to sets J_p and J_g. Inputs of local feed in land use types $j \in J_o$ are defined via coefficients similar to $\bar{\sigma}^f_{j'j}$, but this time expressed per capacity unit of the receiving land use type *j*.

5. Stages in model development and application

This final section reviews the noteworthy features at the various stages of model development and application:

(1) *Base year data*

Chinagro has an extensive database that combines information from a wide range of sources, so as to obtain a comprehensive and consistent classification, while imposing consistency and filling gaps. Data at provincial level play a central role at this stage, even though the province itself is not a geographical unit of the model. Provincial data are aggregated to regional level and serve as benchmark for testing county level information. All economic transactions are expressed in a price and a volume component. Data on agricultural production combine statistical and agronomic sources, while respecting the physical resource balances in each county. The data process for Chinagro 1.0 is described in Van Veen et al. (2005), structured around the derivation of a Chinagro Accounting Matrix for base year 1997. For version 2.0 the process was repeated, this time for base year 2005.

(2) *Model coefficients*

Estimation of coefficients in the functions of the model largely relies on cross-section data, in particular for input relations in agricultural production, supported by technical information on potential yields and resource availability. The parameters of the demand system are based on the 2005 Household Survey of the National Bureau of Statistics of China. Trade and transportation coefficients have been obtained from the price margins in the base year data set. World trade price functions (3.6) have been estimated on the basis of dedicated simulations with the GTAP-model (Hertel, 1997) for a sample of exogenous trade flows of China, supplemented with information on price effects from other worldwide models (AGLINK-COSIMO, FAPRI, IMPACT), which led to upward adjustment of the GTAP-based reaction coefficients.

(3) *Calibration*

The calibration procedure adjusts of a subset of model parameters to ensure that the optimum of model (3.22) will, in base year 2005, at given levels of exogenous variables fully replicate the endogenous variables of the model as recorded in the internally consistent base year data set, also at county level. The calibration can also be looked at as an inclusion of fixed effects in earlier regressions, so as to eliminate regression errors while ensuring that the first-order conditions of the welfare program are met at (and, since the program is strictly convex, only at) a point that agrees with base year data. It proceeds in two phases.

First, the supply calibration reproduces the base year data for farm production and farm input demand in each county at given levels of regional market prices and trade and transportation coefficients between the region and the county, as specified in the 2005 data base. It proceeds in four stages, each with county-specific shifts: (i) calibration of the allocation of area or stable capacity within each land use type, (ii) calibration of the Mitscherlich-Baule input branch for each land use type, (iii) calibration of the piecewise linear input branch for each land use type, and (iv) calibration of the labor allocation among the land use types. The latter is done via imputing a pseudo-output of non-agriculture for each land use type that can be interpreted as reflecting the difference in reward of local deliveries.

Next, at given levels of production and input demand by region, the exchange equilibrium calibration reproduces the base year data for consumer demand, trade flows and market prices at given levels of farm production, farm input demand and processing margins, as specified in the 2005 database. It proceeds in three stages: (i) calibration of consumer demand, (ii) calibration of trade routing, and (iii) calibration of foreign trade and nonagricultural demand.

Simulation models are often validated by showing that its key variables follow an observed pattern. We refer to this as partial calibration, and advocate full calibration as an essential step in model validation. First, it offers a transparent check that the database used to feed the model is internally consistent. Second, since the optimal solution to be found by the optimization algorithm is known beforehand, it also provides a thorough error check on the computer program and the algorithms used for solving this program. Third, as all initial values are known, the calibration also gives an error check for the post-optimal calculations to generate the tables. Finally, and most importantly, it maintains the carefully constructed base year data set as the central empirical basis for the scenario analysis.

After base year calibration, the final stage in validation of the Chinagro model consisted of applying partial validation to the transition from 2005 to 2010.

(4) *Computation*

A model of this size needs a dedicated algorithm. Chinagro is solved via a globally convergent algorithm that decomposes the problem into an exchange component and a supply component coordinated by a price adjustment. The exchange component is an 8-region social welfare maximization that takes output and input of the 2,885 counties as given. This is a regular medium-sized convex program whose optimum is found via a Minos-solve statement in GAMS. The second component, of agricultural supply consists of a series of county-specific farm-income maximization programs that take prices as given and are solved with a tailor-made algorithm that terminates in a finite number of iterations and has an exact solution. This property of finite and exact termination makes it possible to embed both components within a price-adjustment loop (implemented through parameter adjustments in GAMS) and to prove global convergence of the overall process. The algorithm proves to be remarkably fast and precise.[5]

In addition to the finite and exact termination of the supply module, the model owes its computational strength to some further features of the supply-demand decomposition. First, it maintains physical balances and hence feasibility throughout. Second, unlike say, Walrasian tatonnement it inherently keeps all prices non-negative, without need for lower bounds. Third, this price-to-price iteration can be shown to be gradient-related, and consequently, to inherit the convergence property of such a process. For further details see Keyzer and van Veen (2005).

(5) *Scenario simulation*

A model with such a level of detail in classifications is not designed to represent truly endogenous dynamics, as this would inevitably lead to a serious accumulation of prediction errors over time. Therefore, Chinagro simulations assume an exogenous value for a wide range of

[5] Numerical performance is as follows. Starting from given data files and estimates of the consumer demand system, the model calibration and preparation of GAMS-input files for simulation take about 10 minutes, on a regular computer (Intel Core2, Duo CPU, 2 GB RAM, 2 Ghz). A five-period simulation (2005, 2010, 2020, 2030) with tabulation is completed within 30 minutes, at a precision of around 0.15% for the most critical regional commodity price in every year.

driving variables, and provide separate static solutions for the endogenous variables for each year of simulation, given the assumed values of the driving variables. Together, these exogenous variables define a simulation scenario.

Major driving forces are population growth, urbanization and interregional migration, international price developments, changes in farmland and stable capacities, yield improvements in agriculture, non-agricultural growth, trends in food preferences and trade liberalization. The important role played by these driving forces requires us to make a careful and coherent specification of future trends, derived from the literature and our own assessments. Under these assumptions, simulations with the Chinagro model analyze the price-based interaction between the supply behavior of farmers, the demand behavior of consumers and the determination of trade flows by merchants.

The scenarios start from the calibrated base year outcomes and cover the period 2005-2030. Before alternative scenarios can be specified, a baseline scenario is defined. The baseline scenario is not 'business as usual' but should be seen as expected central tendency obtained after ample discussion among the project partners about the assumed trends of the driving forces. In the Chinese context it should even go beyond that and inspire a policy commitment to its realization. This implies that the model outcomes should not only be plausible but also acceptable to policy makers. The most critical outcomes in this respect are the development of farm incomes vis-à-vis the non-farm sector and, in particular, the flows of food and feed that have to be imported from abroad. Comparing the outcomes of these flows with the projections from other models is therefore also part of the exercise. In this respect, we mention in particular the projections of CAPSiM and those reported in OECD-FAO (2011), FAPRI (2011) and USDA (2011).

Model applications so far distinguish six types of scenario, each reflecting specific pathways for the major driving forces: (i) baseline, (ii) trade liberalization, (iii) rapid or slow economic growth, (iv) high agricultural R&D investment, (v) enhanced irrigation efforts, and (vi) various modes of implementation of biofuel policies. For a presentation of earlier results on (i) to (v) we refer to Fischer et al. (2007), whereas biofuel scenarios are discussed in Qiu et al. (2008, 2011). In the companion paper to this one (Keyzer et al., 2012), we discuss the most recent versions of the baseline scenario and the policy variants.

Appendix: Classifications of the Chinagro model

Regions r and their link to provinces

1. North	Beijing, Tianjin, Hebei, Shanxi, Shandong, Henan	(633 counties)
2. Northeast	Liaoning, Jilin, Heilongjiang	(287 counties)
3. East	Shanghai, Jiangsu, Zhejiang, Anhui	(322 counties)
4. Central	Jiangxi, Hubei, Hunan	(324 counties)
5. South	Fujian, Guangdong, Guangxi, Hainan	(340 counties)
6. Southwest	Chongqing, Sichuan, Guizhou, Yunnan	(444 counties)
7. Plateau	Tibet, Qinghai	(119 counties)
8. Northwest	Inner Mongolia, Shaanxi, Gansu, Ningxia, Xinjiang	(416 counties)

This list includes also the distribution of the 2885 counties (defined administratively) over the regions. The regions are depicted in Figure 2.1 of the main text.

Household groups $i \in I_r$ *(in each region)*

1. Rural low income
2. Rural middle income
3. Rural high income
4. Urban low income
5. Urban middle income
6. Urban high income

Tradable commodities k (and volume units)

1. Rice	(kg milled)
2. Wheat	(kg flour)
3. Maize	(kg grain)
4. Other staple food	(kg soybean equivalent)
5. Vegetable oil	(kg)
6. Sugar	(kg refined)
7. Fruits	(kg)
8. Vegetables	(kg)
9. Ruminant meat	(kg carcass weight)
10. Pork	(kg carcass weight)
11. Poultry meat	(kg carcass weight)
12. Milk	(kg fresh milk equivalent)
13. Eggs	(kg)
14. Fish	(kg)
15. Nonfood excl feed	(ten constant 2005 Yuan)
16. Carbohydrate feed	(mcal)
17. Protein feed	(mcal)

Local commodities (and volume units)

1.	Crop residuals	(mcal)
2.	Grass	(mcal)
3.	Green feed	(mcal)
4.	Household waste	(mcal)
5.	Animal manure	(kg nutrient)
6.	Nightsoil	(kg nutrient)
7.	Animal power	(kwh)
8.	Machine power	(kwh)

Feed (combining tradable and local)

Tradable feed:		
1.	Maize	maize grain
2.	Carbohydrate feed	minor and low-quality grain, tubers, vegetables feed, molasses
3.	Protein feed	bran (from wheat and rice), cake (from soybean, peanuts, cottonseed and other oilseeds), fishmeal
Local feed:		
4.	Crop residuals	straw, husk, maize stem and leaves
5.	Grass	fresh grass, hay
6.	Green feed	green fodder, water plants
7.	Household waste	residuals from household consumption

Land use types j and their activities h

1.	Irrigated cropping	All crops
2.	Rainfed cropping	All crops
3.	Tree cropping	Fruits
4.	Draught animal system	Buffaloes, draught cattle, other draught animals
5.	Grazing system	Milk cattle, meat cattle, sheep and goat, yaks
6.	Traditional ruminant farming	Milk cattle, meat cattle, sheep and goat
7.	Specialized dairy farming	Milk cattle
8.	Traditional non-ruminant farming	Hogs, poultry
9.	Intensified non-ruminant farming	Hogs, poultry

Crop activities in this list:

Paddy	Soybeans	Fruits
Wheat	Groundnuts	Vegetables
Maize	Other oil crops	Cotton
Other grains	Sugarcane	Other nonfood crops
Roots and tubers	Sugar beets	

Supply types related to farming

10. Machine power
11. Household waste
12. Household manure

13. Green feed
14. Harvesting grass

Non-farm supply types and their activities

1. Fish and forestry	Fisheries, forestry
2. Non-agriculture	Industry and construction, services

Agricultural resources (with land use or supply type of which they determine the capacity)

1. Irrigated cropland	Irrigated cropping
2. Rainfed cropland	Rainfed cropping
3. Orchard	Tree cropping
4. Natural grassland	Grazing system; harvesting grass
5. Sown grassland	Grazing system; harvesting grass
6. Draught animals	Draught animal system
7. Grazing ruminants	
8. Traditionally mixed ruminants	Traditional ruminant farming
9. Specialized milk cattle	Specialized dairy farming
10. Traditionally mixed non-ruminants	Traditional non-ruminant farming
11. Intensive non-ruminants	Intensified non-ruminant farming
12. Farm labor force	
13. Agricultural machine power	Machine power
14. Greenfeed supply capacity	Green feed
15. Rural population	Household manure; household waste
16. Urban population	Household waste

References

Albersen, P.J., G.F. Fischer, M.A. Keyzer, and L. Sun (2002) Estimation of agricultural production in the LUC-model for China'. IIASA Research Report RR-02-03, International Institute for Applied Systems Research, Laxenburg, Austria.

Arrow, K.J. and F.H. Hahn (1971) *General competitive analysis*. North-Holland.

Brooke, A., D. Kendrick, A. Meeraus, and R. Raman (2011) *GAMS: a user's guide*. GAMS Development Corporation, Washington. Available at http://www.gams.com.

Debreu, G. (1959) *Theory of value: an axiomatic analysis of economic equilibrium*. Cowles Foundation Monograph. New Haven: Yale University Press.

Dervis, K., J. De Melo, and S. Robinson (1982) *General Equilibrium Models for Development Policy*. New York: Cambridge University Press.

Dixon, P., B. Parmenter, G. Ryland and J. Sutton (1977) *Orani : a general equilibrium model for the Australian Economy*. Canberra: Government Printing Office.

FAPRI (2010) US and world agricultural outlook, Staff report 10-FSR 1, Food and Agricultural Policy Research Institute, Iowa State University and University of Missouri-Columbia, Ames.

FAPRI (2011) Characterization of FAPRI models, Food and Agricultural Policy Research Institute, Iowa State University and University of Missouri-Columbia, Ames. Available at http://www.fapri.iastate.edu/models/.

Fischer, G., J. Huang, M.A. Keyzer, H. Qiu, L. Sun and W.C.M. van Veen (2007) China's agricultural prospects and challenges: report on scenario simulations with the Chinagro welfare model covering national, regional and county level. Amsterdam: Centre for World Food Studies.

Ginsburgh, V. and M.A. Keyzer (2002) *The structure of applied general equilibrium models*. Cambridge MA: MIT Press.

Gunning, J.W., and M.A. Keyzer (1995) Appled General Equilibrium Models for Policy Analysis. In J. Behrman and T.N. Srinivasan, eds. *Handbook of Development Econonomics*, Vol. 3. Amsterdam: North Holland.

Hertel, T.W., ed. (1997) *Global trade analysis: modeling and applications*. Cambridge: Cambridge University Press.

Huang, J. and Li, N. (2003) 'China's agricultural policy simulation and projection model: CAPSiM', *Journal of Nanjing Agricultural University (Social Science Edition)*, 3: 30-41.

Keyzer, M.A. and W.C.M. van Veen (2005) Towards a spatially and socially explicit agricultural policy analysis for China: specification of the Chinagro models, Working Paper 05-02, Amsterdam: Centre for World Food Studies.

Keyzer, M.A., W.C.M. van Veen, J. Huang, H. Qiu, T. Ermolieva and G. Fischer (2012) China's food demand, supply and trade in 2030: simulations using the Chinagro model, CATSEI paper, Amsterdam: Centre for World Food Studies, forthcoming.

OECD (2007) Documentation of the AGLINK-COSIMO model, Report AGR/CA/APM (2006)16, Committee for Agriculture, Organisation for Economic Co-operation and Development, Paris.

OECD-FAO (2011) Agricultural outlook 2011-2020, Organisation for Economic Co-operation and Development and Food and Agriculture Organization of the United Nations, Paris/Rome.

Rosegrant, M.W., C.Ringler, S. Msangi, T.B. Sulser, T. Zhu and S.A. Cline (2008) International Model for Policy Analysis of Agricultural Commodities and Trade (IMPACT): model description, International Food Policy Research Institute, Washington.

Qiu, H., J. Huang, M.A. Keyzer and W.C.M. van Veen (2008) 'Policy options for China's bio-ethanol development and the implications for its agricultural economy', *China & World Economy* 16: 112-124.

Qiu, H., J. Huang, M.A. Keyzer, W.C.M. van Veen, S. Rozelle, G. Fischer and T. Ermolieva (2011) 'Biofuel development, food security and the use of marginal land in China', *Journal of Environmental Quality* 40: 1058-1067.

USDA (2011), Agricultural projections to 2020, Interagency Agricultural Projections Committee, United States Department of Agriculture, Washington.

Veen, W.C.M. van, P. J. Albersen, G. Fischer and L. Sun (2005) Data set of the Chinagro welfare model: structure and composition, Working Paper 05-03. Amsterdam: Centre for World Food Studies.

The Centre for World Food Studies (Dutch acronym SOW-VU) is a research institute related to the Department of Economics and Econometrics of the Vrije Universiteit Amsterdam. It was established in 1977 and engages in quantitative analyses to support national and international policy formulation in the areas of food, agriculture and development cooperation.

SOW-VU's research is directed towards the theoretical and empirical assessment of the mechanisms which determine food production, food consumption and nutritional status. Its main activities concern the design and application of regional and national models which put special emphasis on the food and agricultural sector. An analysis of the behaviour and options of socio-economic groups, including their response to price and investment policies and to externally induced changes, can contribute to the evaluation of alternative development strategies.

SOW-VU emphasizes the need to collaborate with local researchers and policy makers and to increase their planning capacity.

SOW-VU's research record consists of a series of staff working papers (for mainly internal use), research memoranda (refereed) and research reports (refereed, prepared through team work).

Centre for World Food Studies
SOW-VU
De Boelelaan 1105
1081 HV Amsterdam
The Netherlands

Telephone (31) 20 – 598 9321
Telefax (31) 20 – 598 9325
Email pm@sow.vu.nl
www http://www.sow.vu.nl/