



FREDERICK S. PARDEE CENTER
FOR INTERNATIONAL FUTURES
EXPLORE UNDERSTAND SHAPE

WORKING PAPER 2015.07.20

IFs ECONOMIC MODEL DOCUMENTATION

Author: Barry B. Hughes
July 2015

Note: If cited or quoted, please indicate working paper status.



UNIVERSITY *of*
DENVER

JOSEF KORBEL SCHOOL OF INTERNATIONAL STUDIES

IFs Economic Model Documentation

Table of Contents

1.	Introduction.....	4
1.1	Overview.....	4
1.2	Dominant Relations	4
1.3	Structure and Agent System.....	5
1.4	Social Accounting Matrix Approach in IFs	7
2.	Flow Charts.....	10
2.1	Economics Overview	10
2.2	Value Added	12
2.3	Multifactor Productivity.....	13
2.4	Economic Aggregates and Indicators	14
2.5	Household Accounts	15
2.6	Firm Accounts.....	16
2.7	Government Accounts	17
2.8	Savings and Investment	19
2.9	Trade	20
2.10	International Finance	21
2.11	Income Distribution	23
2.12	Poverty	24
2.13	Informal Economy	25
3.	Equations.....	27
3.1	Overview.....	27
3.2	The Goods and Services Market	27
3.2.1	<i>The Production Function: Basic Overview</i>	27
3.2.2	<i>The Production Function: Detail</i>	30
3.2.3	<i>Gross Production and Intersectoral Flows</i>	38
3.2.4	<i>Labor Supply</i>	41
3.2.5	<i>Capital-Labor Shares</i>	44
3.2.6	<i>GDP at MER and PPP and Relative-price Adjusted GDP</i>	44
3.2.7	<i>Trade</i>	45
3.2.8	<i>Stocks and Prices</i>	49
3.2.9	<i>Investment</i>	49
3.2.10	<i>Indicators</i>	51
3.3	The Social Accounting Matrix (SAM)	51
3.3.1	<i>Data for Initial Conditions in the Pre-processor</i>	51
3.3.1.1	A note on forecasting central and local government finance.....	52
3.3.1.2	Possible futures steps in forecasting central/local government finance....	53
3.3.2	<i>Government Revenue</i>	54
3.3.3	<i>Government Expenditure</i>	55
3.3.4	<i>Government Consumption by Destination</i>	60
3.3.5	<i>Government Balances and Dynamics</i>	67

3.3.6	Household and Firm Income	69
3.3.7	Household Consumption and Net Savings	71
3.3.7.1	The permanent income foundation of consumption	72
3.3.7.2	Basic consumption computation and the affect of GDP per capita	72
3.3.7.3	The impact of age structure on consumption	73
3.3.7.4	The impact of life expectancy on consumption	75
3.3.7.5	Interest rates, stocks (inventories) and consumption	76
3.3.7.6	Savings rates and consumption	77
3.3.7.7	Investment and consumption	77
3.3.7.8	Consumption by household type and update of consumption propensity for future years	78
3.3.7.9	Household savings	79
3.3.8	Household Consumption by Sector	79
3.3.9	Domestic Distribution of Household Income	80
3.3.9.1	Conceptual foundations of distribution in IFs: The Lorenz Curve and Gini Coefficient	80
3.3.9.2	An earlier, legacy process to compute domestic Gini	83
3.3.9.3	The current process in IFs to compute domestic and global Gini coefficients	85
3.3.10	Firm Savings	87
3.3.11	Domestic Closure: Total Savings, Investment, and Interest Rates	87
3.3.11.1	Foreign and total savings	87
3.3.11.2	Gross capital formation and investment	89
3.3.11.3	Interest rates	90
3.3.12	Current and Capital Account Elements	91
3.3.12.1	Foreign Aid	92
3.3.12.2	Worker remittances	93
3.3.12.3	Foreign Direct Investment	94
3.3.12.4	Portfolio Flows	97
3.3.12.5	International financial institution flows	100
3.3.13	International Accounting and Financial Equilibration	104
3.3.13.1	Current Account	105
3.3.13.2	Capital Account	106
3.3.13.3	External Debt	106
3.3.13.4	External Debt for exchange rates and external accounts balancing	107
3.3.13.5	Foundations in the preprocessor for extending treatment of debt	108
3.4	Poverty	109
3.4.1	The Proximate Drivers of Poverty	109
3.4.2	Lognormal Analysis of Change in Poverty: IFs Formulations	114
3.4.3	Cross-sectional Analysis of Change in Poverty	116
3.5	Informal Economy	117
3.5.1	Drivers of Informality	118
3.5.2	Forecasting Informal Labor Share	118
3.5.3	Forecasting Informal Labor by Sector	120
3.5.4	Forecasting Informal GDP Share	121
3.5.4.1	Shadow economy	123

3.5.5	<i>Forward Linkages From Informal Economy</i>	123
3.5.5.1	Forward linkages: productivity and GDP	124
3.5.5.2	Forward linkages: government revenue	127
3.6	The Adjustment Mechanism and Other Specialized Functions	128
3.7	Creating and Initializing Country SAMS: The Pre-processor	129
4.	References	131

1. Introduction

1.1 Overview

The economics model of IFs forecasting system draws on two general modeling traditions. The first is the dynamic growth model of classical economics. Within IFs the growth rates of labor force, capital stock, and multifactor productivity largely determine the overall size of production and therefore of the economy. The second tradition is the general equilibrium model of neo-classical economics. IFs contains a six-sector (agriculture, raw materials, energy, manufactures, services, and ICT) equilibrium-seeking representation of domestic supply, domestic demand, and trade. Further, the goods and services market representation is embedded in a larger social accounting matrix structure that introduces the behavior of household, firm, and government agent classes and the financial flows they determine.

1.2 Dominant Relations

In any long-term economic model the supply side has particular importance. In IFs, gross domestic product (GDP) is a function of multifactor productivity (MFP), capital stocks (KS), and labor inputs (LABS), all specified for each of six sectors. This approach is sometimes called a Solowian Cobb-Douglas specification, but IFs helps the user get inside the multifactor productivity term, rather than leaving it as a totally exogenous residual.

The following key dynamics are directly related to the dominant relations:

- Multifactor productivity is a function partly of exogenous specification of an annual growth rate in it for the systemic technology leader, base rates of relative technological advance in other countries determined via an inverted U-shaped function that assumes convergence with the leader, and of an exogenously specified additive factor for control of specific regions or countries.
- Multifactor productivity is, however, largely an endogenous function of variables determined in other models of the IFs system representing the extent of human, social, physical, and knowledge capital; their influence on production involves coefficients that the user can control.
- Capital stock is a function of investment and depreciation rates. Endogenously determined investment can be influenced exogenously by a multiplier and the lifetime of capital can be changed.
- Labor supply is determined from population of appropriate age in the population model (see its dominant relations and dynamics) and endogenous labor force participation rates, influenced exogenously by the growth of female participation.

The larger economic model provides also representation of and some control over sector-specific consumption patterns; trade including protectionism levels and terms of trade; taxation levels; economic freedom levels; and financial dynamics around foreign aid, borrowing, and external debt.

1.3 Structure and Agent System

Goods and Services Market

System/Subsystem	Goods and Services
Organizing Structure	Endogenously driven production function represented within a dynamic general equilibrium-seeking model
Stocks	Capital, labor, accumulated technology
Flows	Production, consumption, trade, investment
Key Aggregate Relationships (illustrative, not comprehensive)	Production function with endogenous technological change; price movements equilibrate markets over time
Key Agent-Class Behavior Relationships (illustrative, not comprehensive)	Households and work/leisure, consumption, and female participation patterns; Firms and investment; Government decisions on revenues and on both direct expenditures and transfer payments

Households, firms, and the government interact via markets in goods and services. There are obvious stock and flow components of markets that are desirable and infrequently changed in model representation. Perhaps the most important key aggregate relationship is the production function. Although the firm is an implicit agent-class in that function, the relationships of production even to capital and labor inputs, much less to the variety of technological and social and human capital elements that enter a specification of

endogenous productivity change (Solow 1957; Romer 1994), involve multiple agent-classes. In the representation of the market now in IFs there are also many key agent-class relationships as suggested by the table.

Financial Flows/Social Accounting

System/Subsystem	Financial
Organizing Structure	Market plus socio-political transfers in Social Accounting Matrix (SAM)
Stocks	Government, firm, household assets/debts
Flows	Savings, consumption, FDI, foreign Aid, IFI credits/grants, government expenditures (military, health, education, other) and transfers (pensions and social transfers)
Key Aggregate Relationships (illustrative, not comprehensive)	Exchange rate, movements with net asset/current account level; interest rate movements with savings and investment
Key Agent-Class Behavior Relationships (illustrative, not comprehensive)	Household savings/consumption; Firm investment/profit returns and FDI decisions; Government revenue, expenditure/transfer payments; IFI credits and grants

Households, firms, and the government interact in markets, but more broadly also via financial flows, including those related to the market (like foreign direct investment), but

extending also to those that have a socio-political basis (like government to household transfers). A key structural representation is the Social Accounting Matrix (SAM).

The structural system portrayed by SAMs is well represented by stocks, flows, and key relationships. Although the traditional SAM matrix itself is a flow matrix, IFs has introduced a parallel stock matrix that captures the accumulation of assets and liabilities across various agent-classes. The dynamic elements that determine the flows within the SAM involve key relationships, such as that which constrains government spending or forces increased revenue raising when government indebtedness rises. Many of these, as indicated in the table, represent agent-class behavior.

The model can represent the behavior of households with respect to use of time for employment and leisure, the use of income for consumption and savings, and the specifics of consumption decisions across possible goods and services. And it represents the behavior of governments with respect to search for income and targeting of transfers and expenditures, in interaction with other agents including households, firms, and international financial institutions (IFIs).

IFs thus represents equilibrating markets (domestically and globally) in goods and services and in financial flows. It does not yet include labor market equilibration,

1.4 Social Accounting Matrix Approach in IFs

A SAM integrates a multi-sector input-output representation of an economy with the broader system of national accounts, also critically representing flows of funds among societal agents/institutions and the balance of payments with the outside world. Richard Stone is the acknowledged father of social accounting matrices, which emerged from his participation in setting up the first systems of national accounts or SNA (see Pesaran and Harcourt 1999 on Stone's work and Stone 1986). Many others have pushed the concepts and use of SAMs forward, including Pyatt (Pyatt and Round 1985) and Thorbecke (2001). So, too, have many who have extended the use of SAMs into new frontiers. One such frontier is the additional representation of environmental inputs and outputs and the creation of what are coming to be known as social and environmental accounting matrices or SEAMs (see Pan 2000). Another very productive extension is into the connection between SAMs and technological systems of a society (see Khan 1998; Duchin 1999). It is fitting that the 1993 revision of the System of National Accounts by the United Nations began explicitly to move the SNA into the world of SAMs.

The SAM of IFs is integrated with a dynamic general equilibrium-seeking model. The structural representation is a variant and to some degree an extension of the computable general equilibrium formulations that often surround SAMs. In wrapping SAMs into CGEs, Stone was a pioneer, leading the Cambridge Growth Project with Alan Brown. That project placed SAMs into a broader modeling framework so that the effects of changes in assumptions and coefficients could be analyzed, the predecessor to the development and use of computable general equilibrium (CGE) models by the World Bank and others. Some of the Stone work continued with the evolution of the Cambridge Growth Model of the British economy (Barker and Peterson, 1987). Kehoe (1996)

reviewed the emergence of applied general equilibrium (GE) models and their transformation from tools used to solve for equilibrium under changing assumptions at a single point in time to tools used for more dynamic analysis of societies.

The approach of IFs is both within these developing traditions and an extension of them on five fronts. The first extension is in universality of the SAM representation. Most SAMs are for a single country or a small number of countries or regions within them (e.g., see Bussolo, Chemingui, and O'Connor 2002 for a multi-regional Indian SAM within a CGE). The IFs project has created a procedure for constructing relatively highly aggregated SAMs from available data for all of the countries it represents, relying upon estimated relationships to fill sometimes extensive holes in the available data. Jansen and Vos (1997: 400-416) refer to such aggregated systems as using a "Macroeconomic social Accounting Framework." Each SAM has an identical structure and they can therefore be easily compared or even aggregated (for regions of the world).

The second extension is the connecting of the universal set of SAMs through representation of the global financial system. Most SAMs treat the rest of the world as a residual category, unconnected to anything else. Because IFs contains SAMs for all countries, it is important that the rest-of-the-world categories are mutually consistent. Thus exports and imports, foreign direct investment inflows and outflows, government borrowing and lending, and many other inter-country flows must be balanced and consistent.

The third extension is a representation of stocks as well as flows. Both domestically and internationally, many flows are related to stocks. For instance, foreign direct investment inflows augment or reduce stocks of existing investment. Representing these stocks is very important from the point of view of understanding long-term dynamics of the system because those stocks, like stocks of government debt, portfolio investment, IMF credits, World Bank loans, reserve holdings, and domestic capital stock invested in various sectors, generate flows that affect the future. Specifically, the stocks of assets and liabilities will help drive the behavior of agent classes in shaping the flow matrix.

The IFs stock framework has been developed with the asset-liability concept of standard accounting method. The stock framework is also an extension of the social accounting flow matrix, and the cumulative flows over time among the agents will determine the stocks of assets or liabilities for all agents. If the inflow demands repayment or return at some point in future, it is considered as liability for that agent and an asset for the agent from which the flow came. For example, in IFs, if a government receives loans (inflow) from other countries, the stock of those loans is a liability for the recipient government and an asset for the country or countries providing the loans.

The fourth extension is temporal and builds on the third. The SAM structure described here has been embedded within a long-term global model. The economic module of IFs has many of the characteristics of a typical CGE, but the representation of stocks and related agent-class driven behavior in a consciously long-term structure introduces a quite different approach to dynamics. Instead of elasticities or multipliers on various terms in the SAM, IFs seeks to build agent-class behavior that often is algorithmic rather than

automatic. To clarify this distinction with an example, instead of representing a fixed set of coefficients that determine how an infusion of additional resources to a government would be spent, IFs increasingly attempts partially to endogenize such coefficients, linking them to such longer-term dynamics as those around levels of government debt. Similarly, the World Bank as an actor or agent could base decisions about lending on a wide range of factors including subscriptions by donor states to the Bank, development level of recipients, governance capacity of recipients, existing outstanding loans, debt-to-export ratios, etc. Much of this kind of representation is in very basic form at this level of development, but the foundation is in place.

The fifth and final extension has already been discussed. In addition to the SAM, The IFs forecasting system also includes a number of other models relevant to the analysis of longer-term forecasts. For example demographic, education, health, agriculture, and energy models all provide inputs to the economic model and SAM, as well as responding to behavior within it. The effort is to provide a dynamic base for forecasts can be made well into the 21st century. It is important to quickly emphasize that such forecasts are not predictions. Instead they are scenarios to be used for thinking about possible alternative longer-term futures.¹

¹ As a graduate student in what is now the Josef Korbel School of International Studies, Anwar Hossain worked with Barry Hughes in the development of the SAM structure and database for IFs (see Hughes and Hossain 2003); his help was much appreciated.

2. Flow Charts

This section presents several block diagrams that are central to the two major components of the economics model, the goods and services market—with special emphasis on the production function—and the broader SAM.

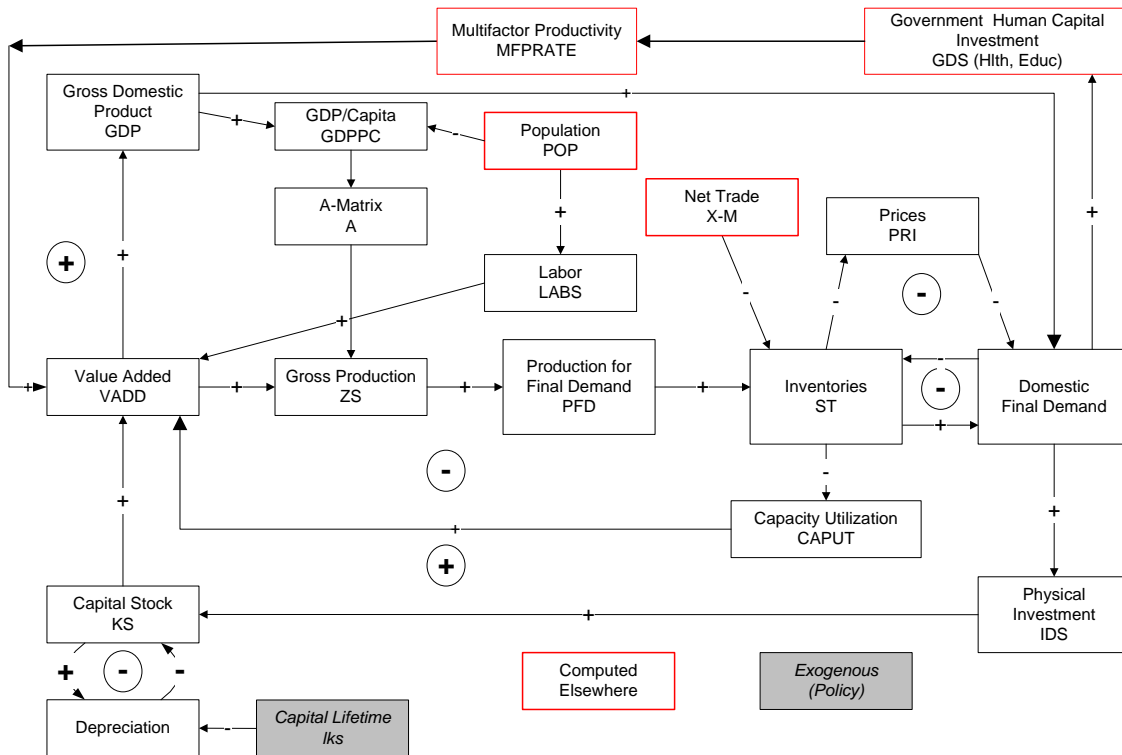
2.1 Economics Overview

The economic model represents supply, demand, and trade in each of six economic sectors: agriculture, primary energy, raw materials, manufactures, services, and information/communications technology. The model draws upon data from the Global Trade and Analysis Project (GTAP) with 57 sectors as of GTAP 8; the pre-processor collapses those into the six IFs sectors and theoretically could collapse them into a different aggregated subset.

Inventories (or stocks) are the key equilibrating variable in three negative feedback loops. As they rise, prices fall, increasing final demand (one loop), decreasing production (a second loop), and thereby in total decreasing inventories in the pursuit over time of a target value and equilibrium. Similarly, as inventories rise, capacity utilization falls, decreasing production, and restraining inventories.

Physical investment and capital stocks are the key driving variables in an important positive feedback loop. As capital rises, it increases value added and GDP, increasing final demand and further increasing investment. Similarly, government social investment can increase productivity, production and inventories in another positive feedback loop.

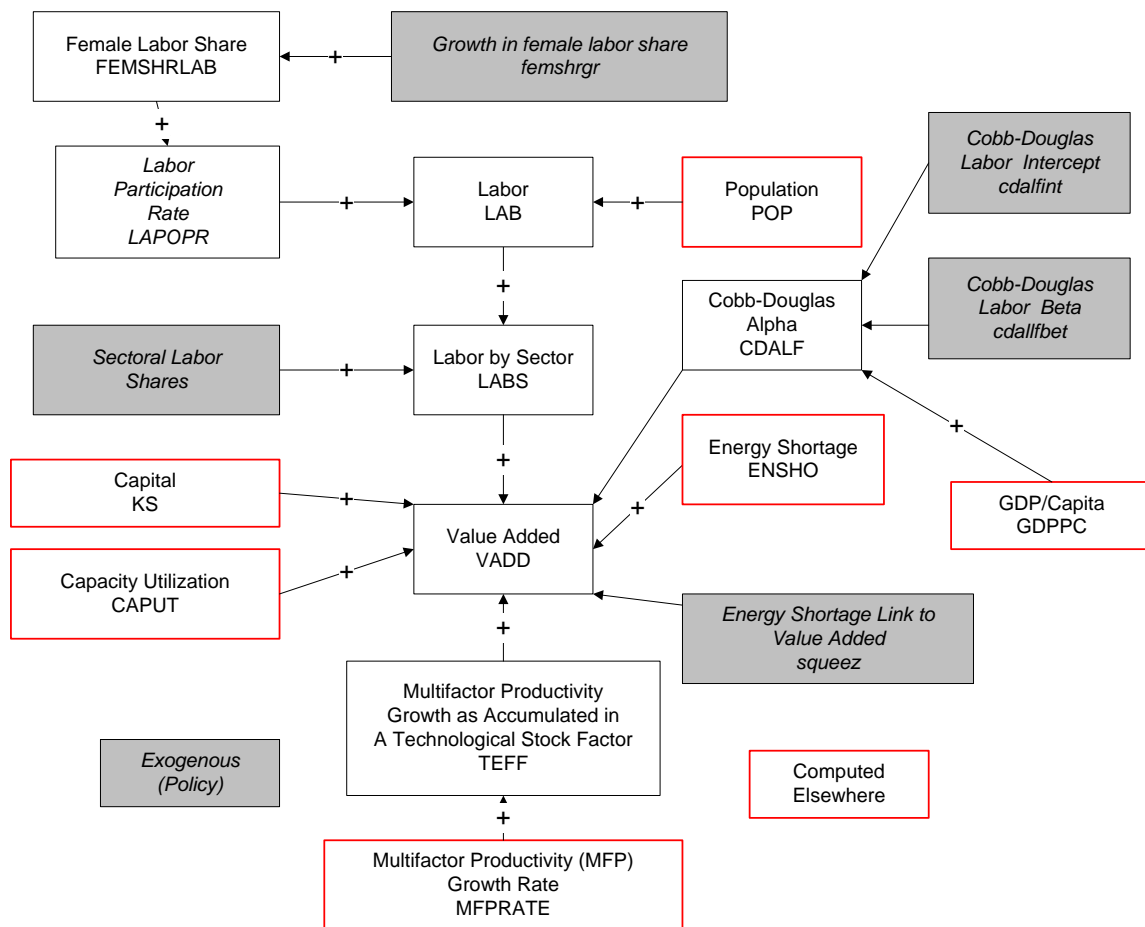
The figure below also shows some production detail. A-matrices, which are specified dependent on the level of development (GDP per capita), allow the computation from value added of gross production and of the production that is available, after satisfaction of intersectoral flows, for meeting final demand. It is the balance of this production for final demand with actual final demand that determines whether inventories grow or decline.



The calculation of gross production (ZS) in value terms within the economic model is overridden by calculations of physical production converted to value in the agricultural and energy models when respective switches (AGON and ENON) are thrown as in the default of the IFs Base Case scenario.

2.2 Value Added

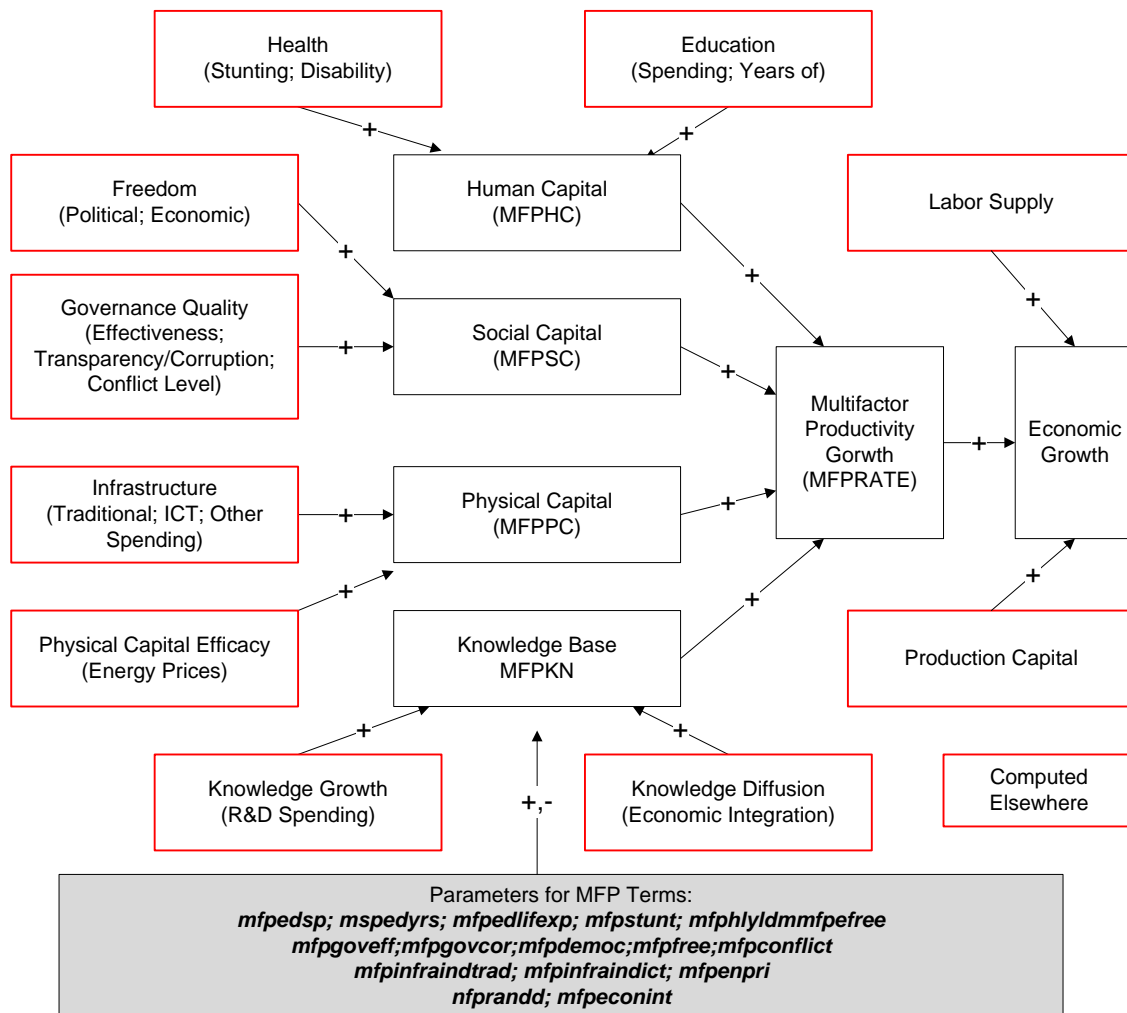
A Cobb-Douglas production function determines value added. Thus two principal factors are capital and labor. Labor is responsive not just to population size and structure, but to the labor participation rate, including the changing role of women in the work force. Accumulated growth in the level of technology or multifactor productivity (MFP), in a "disembodied" representation (TEFF), modifies these factors. Immediate energy shortages/shocks can also affect value added.



2.3 Multifactor Productivity

The technological factor in the production function is often called multifactor productivity (MFP). The basic value of MFP is a sum of a global productivity growth rate driven by the economically advanced or leading country/region (*mfpleadr*), a technological premium dependent on GDP per capita, and an exogenous or scenario factor (*mfpadd*).

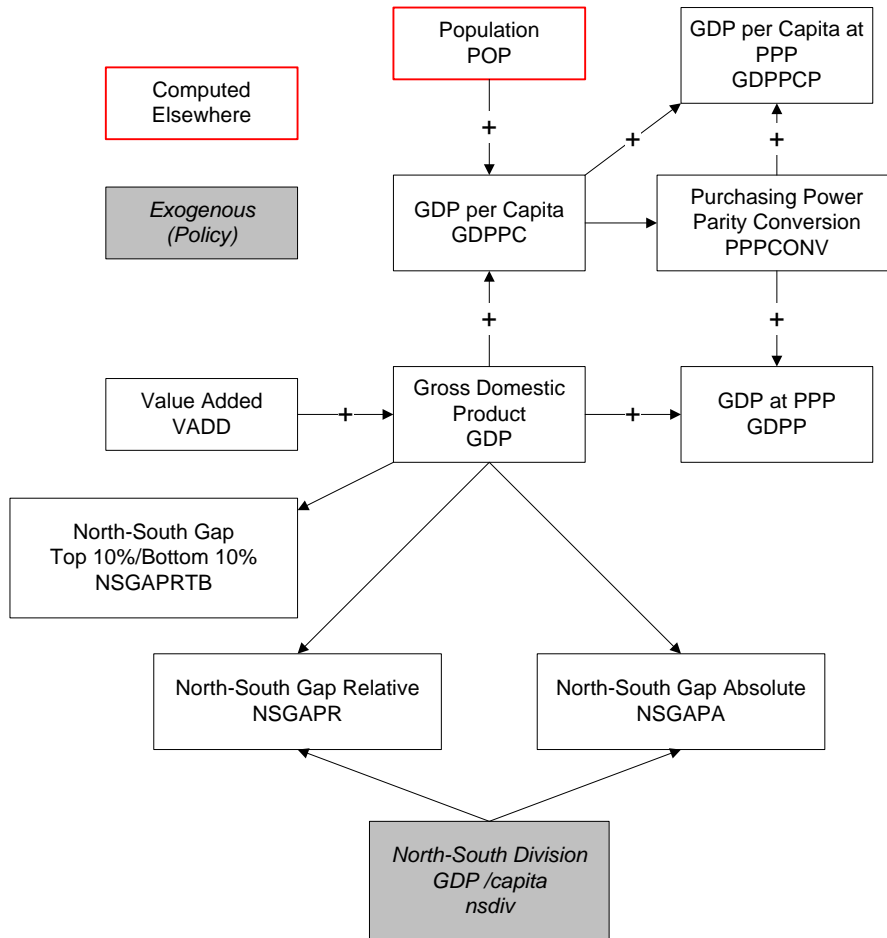
In addition, however, other factors affect productivity growth over time. These include a wide range of variables, such as the years of education that adults have (EDYRSAG25) and the level of economic freedom (ECONFREE), which respectively are among the variables that affect change in MFP associated with human and social capital.



*

2.4 Economic Aggregates and Indicators

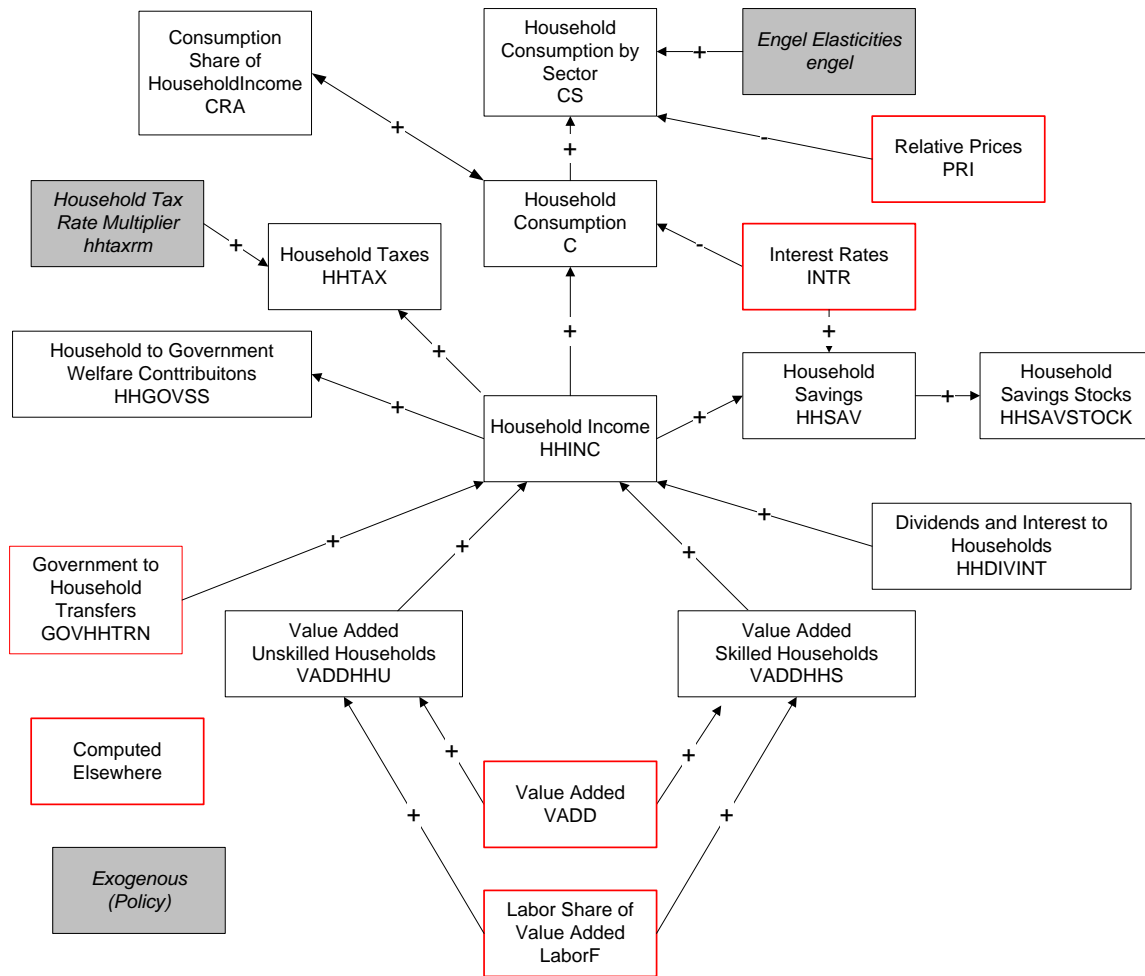
Based on value added and population, it is possible to compute GDP, GDP at purchasing power parity, and a substantial number of country/region-specific and global indicators including several that show the extent of the global North/South gap.



2.5 Household Accounts

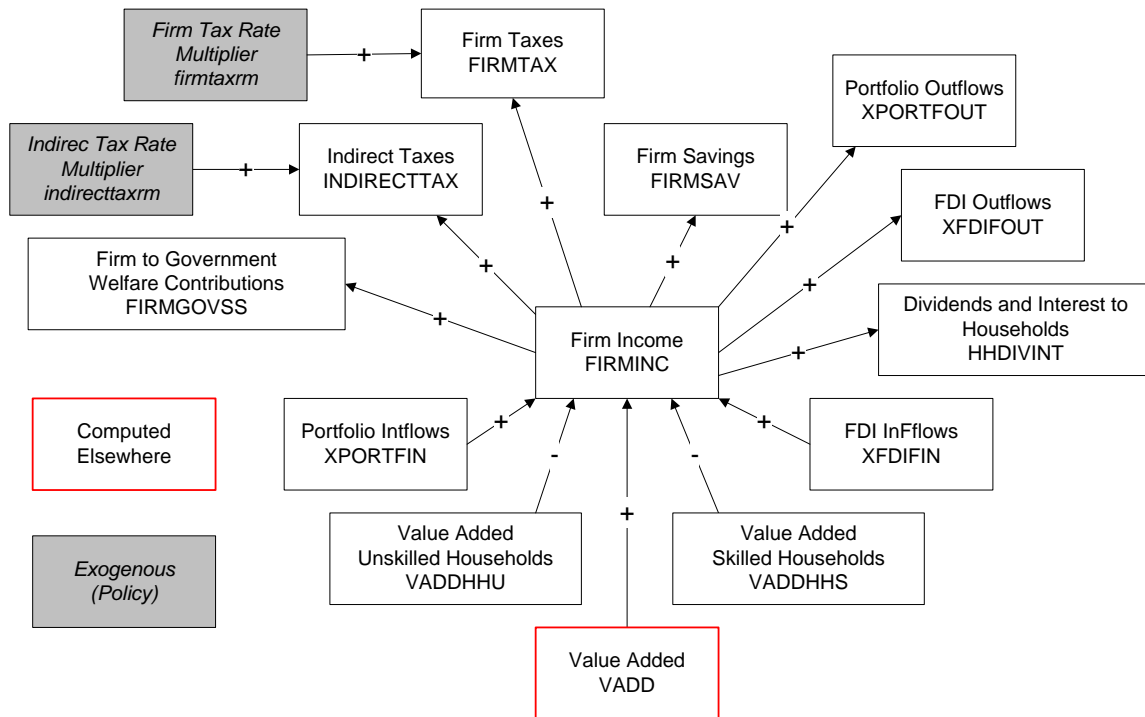
The most important drivers of household income is the size of value added and the share of that accruing to households. That share, divided further into unskilled and skilled households, is initialized with data from the Global Trade and Analysis Project and changes a function driven by GDP per capita. Household income is augmented by flows from government and firms (dividends and interest). Most of household income will be used for consumption, but shares will go back to the government via taxes and to savings.

Once the total of household consumption is known, it is divided across the sectors of IFs using Engel elasticities that recognize changing use of consumption as levels per capital rise.



2.6 Firm Accounts

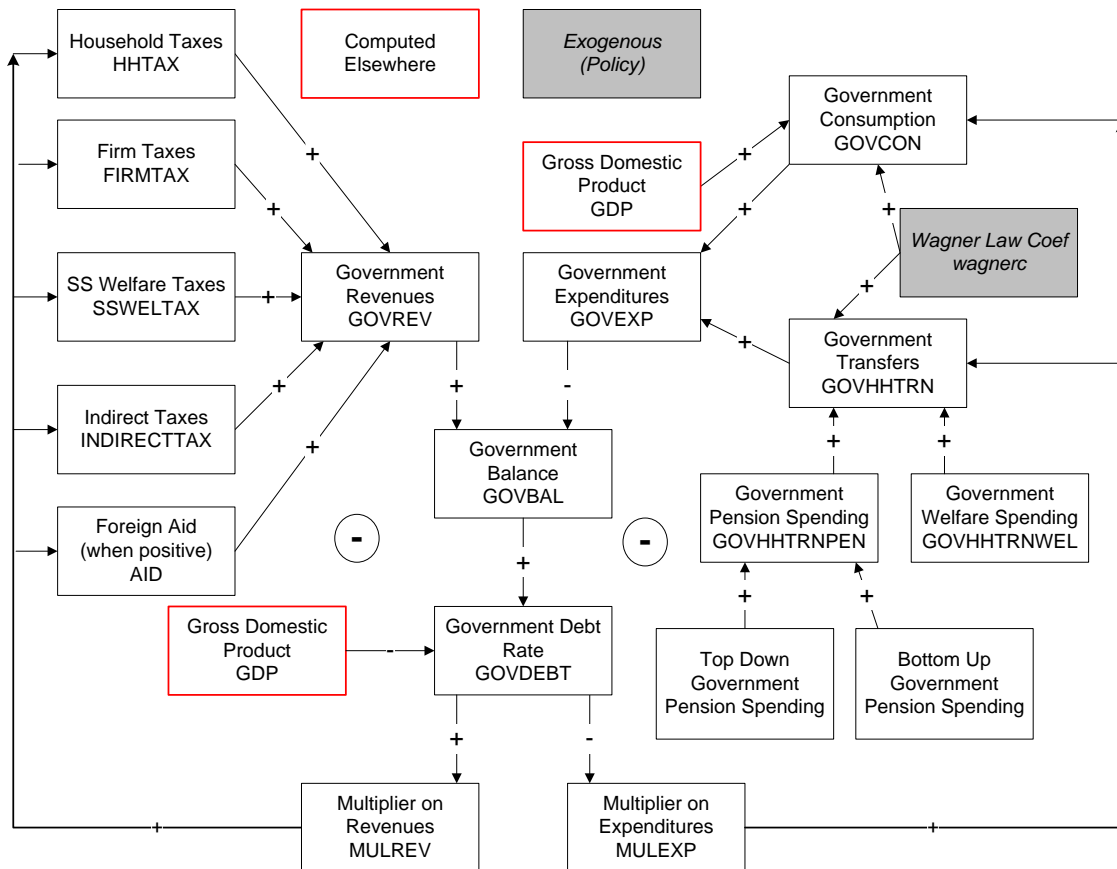
Firms retain as income the portion of total value added that is not sent to households in return for labor provided. Income of firms functioning within a country (ownership in IFs is not designated as domestic or international) benefit from inflows of portfolio and foreign direct investment. Firms direct their income to governments in the form of tax payments, to households as dividends and interest, to the outside world as portfolio or foreign direct investment, and to savings (available for investment).



2.7 Government Accounts

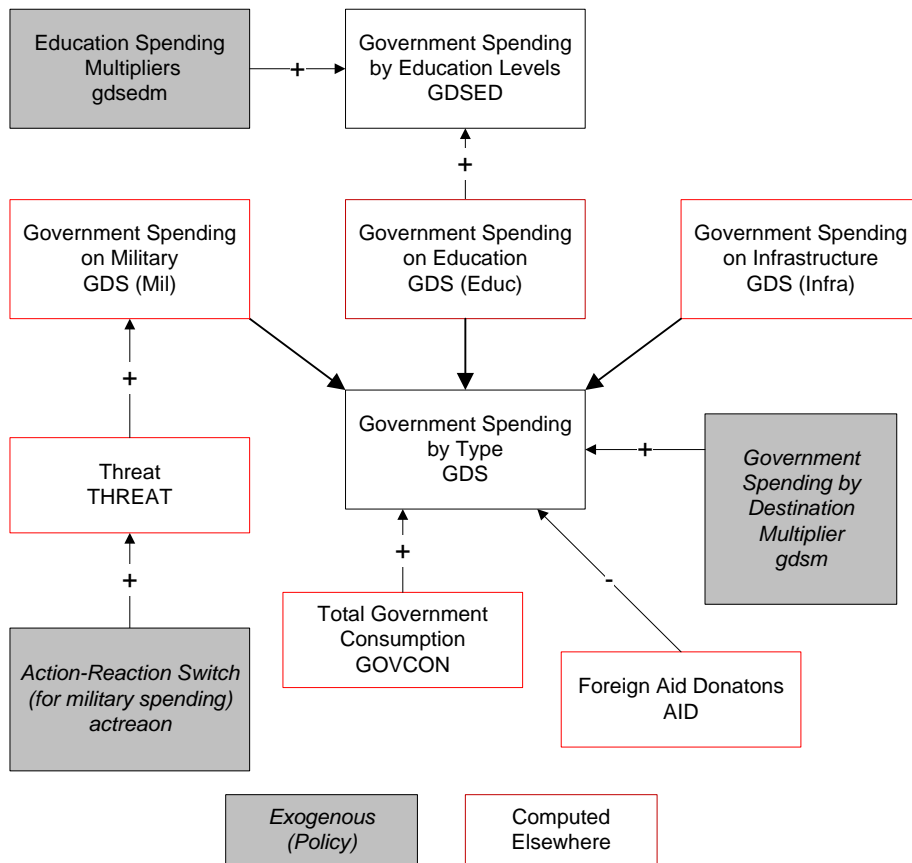
Government revenues come from taxes levied on households and firms. The total expenditures are a sum of two sub-categories, direct consumption and transfer payments (the latter in turn being a sum of payments for pensions/retirement and welfare).

The annual government balance is the difference between revenues and expenditures and increments or decrements government debt in absolute terms and as a portion of GDP. That stock variable in turn sends back signals to both revenue and expenditure sides of the model so as to keep the debt at reasonable levels over time.



The level of government consumption and its distribution across targets are important policy-relevant variables in the model. Government consumption is spread across several target spending categories: military health, education, traditional infrastructure, other infrastructure, research and development, other/residual, and foreign aid). The distribution to most of those categories is endogenously determined by functions, but other models in the IFs system provide special signals for military, education, and traditional infrastructure spending. Demand for military spending involves action-reaction dynamics (when a switch is turned on) and threat levels. Demand for educational and infrastructure involves full models. Demands will not equal supply and all demands are normalized to the supply, but special protection can be given to the demands for education and infrastructure spending.

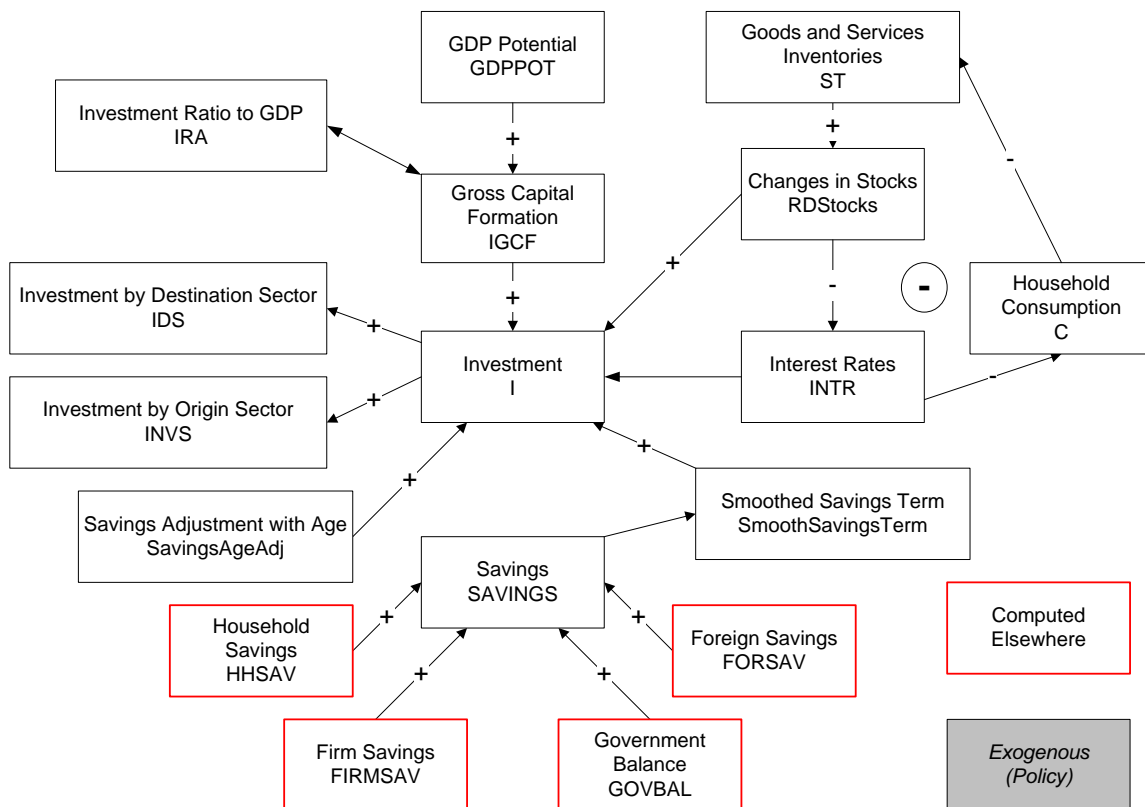
Educational spending by level of education (primary, secondary, and tertiary) is further broken out of total educational spending in the education model but targets can be changed via a multiplier.



2.8 Savings and Investment

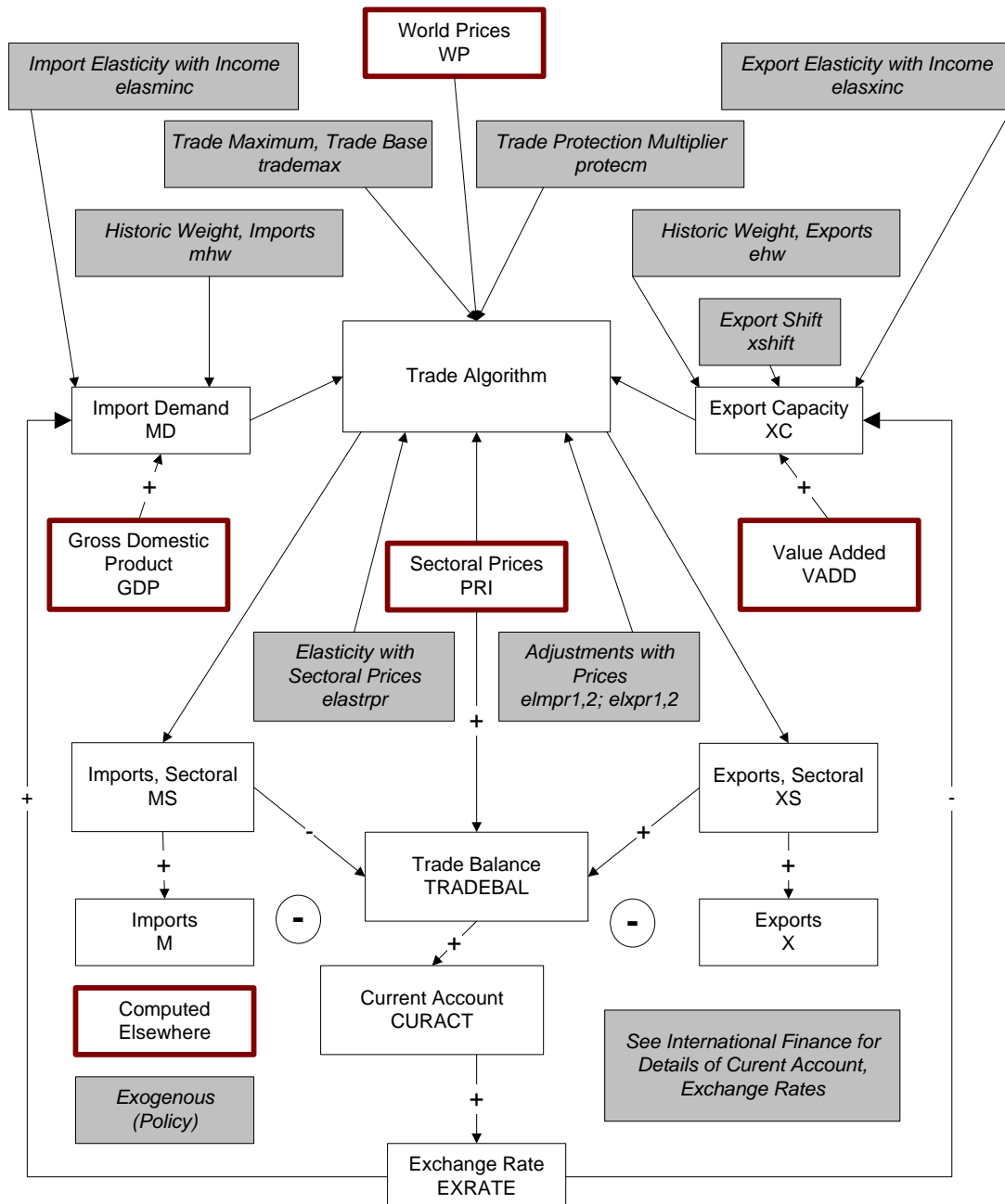
Savings is a sum of the savings by households, firms, the government (its fiscal balance) and net foreign savings. Investment is most immediately a sum of gross capital formation and changes in inventories.

As in other parts of the IFs economic model, there will not be an exact equilibrium between savings and investment in any given time step. The system will chase equilibrium over time with the help of two mechanisms. The smoothed pattern of savings over time will affect investment. So, too, will interest rates that respond to changes in inventories or stocks of goods and services.



2.9 Trade

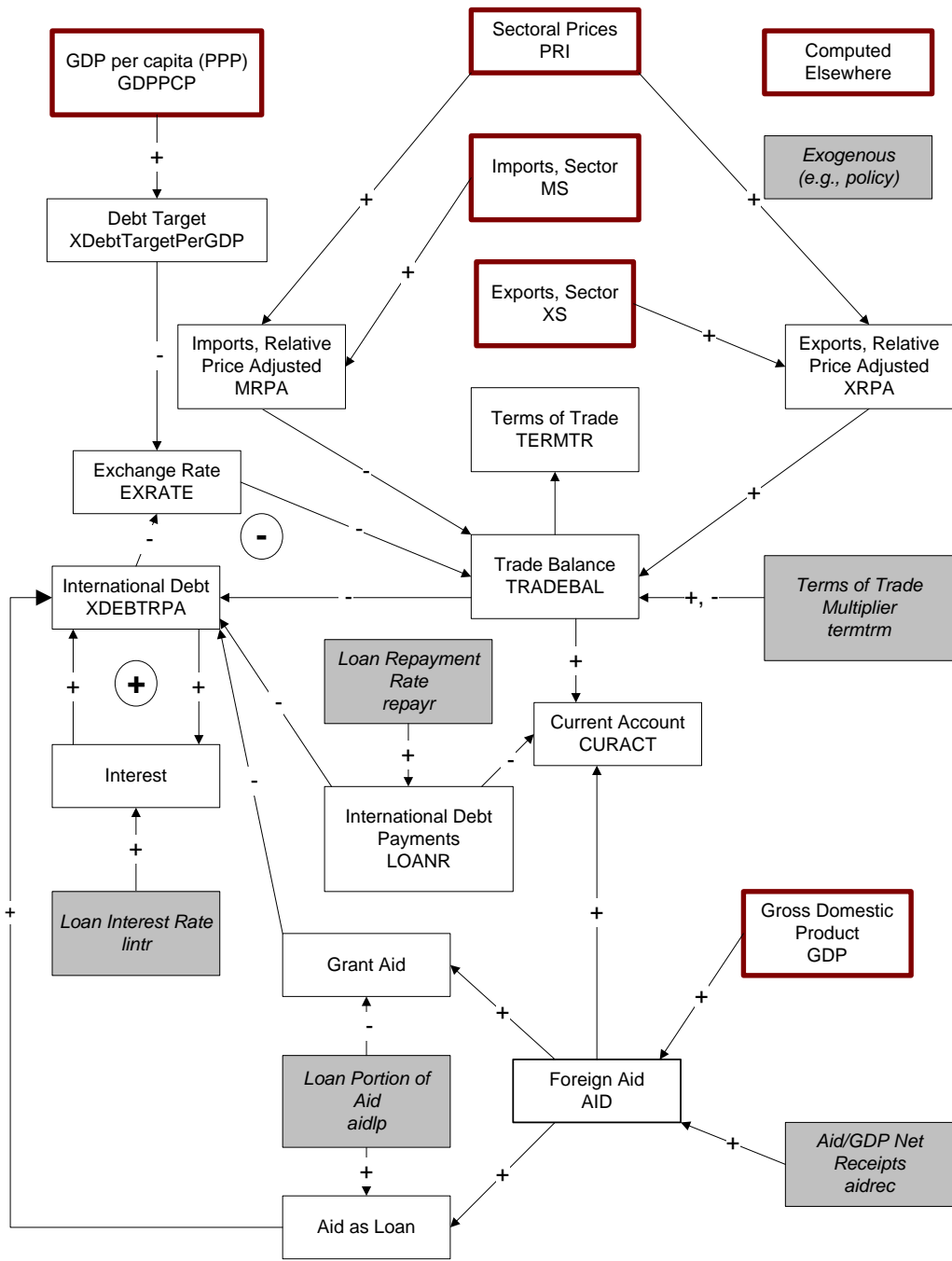
The trade algorithm of IFs relies on a pooled rather than bilateral trade approach. That is, it does not track exactly who trades with whom, maintaining instead information on gross exports and imports by sector and in total. The algorithm sums import demand and export capacity across all traders (in a given sector), defines world trade as the average of those two values, and then normalizes demand and capacity to the total of world trade to determine sectoral exports and imports by geographic unit.



World trade as a percentage of GDP (WTRADE) is also computed from the trade variables shown above.

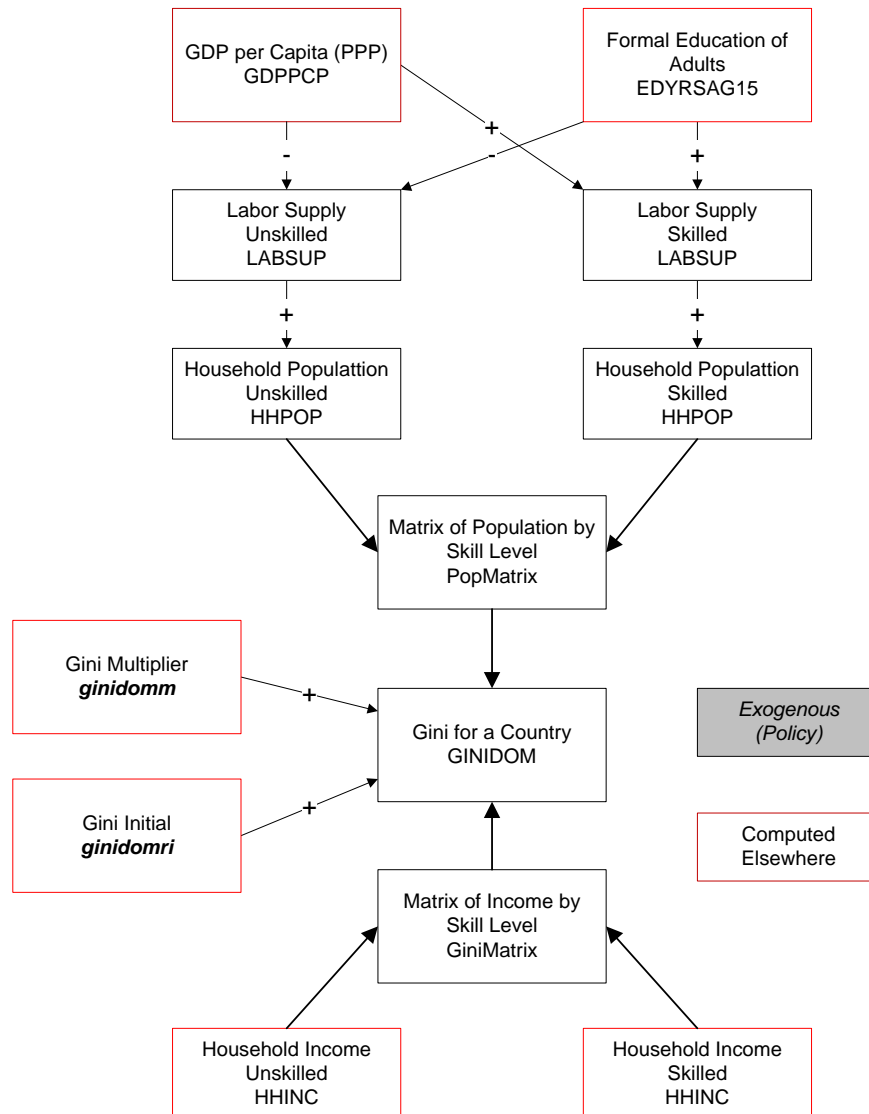
2.10 International Finance

The current account depends on international loan repayments and foreign aid flows as well as on the trade balance. The exchange rate floats with the debt level (in turn responsive to the current account balance) and is the key equilibrating variable in two negative feedback loops that work via import demand and export capacity (see description of trade).



2.11 Income Distribution

Domestic income distribution is represented by the Gini coefficient. That is calculated with an Lorenz curve that looks at the share of population and income held by the only two subgroups for which we have that information, namely unskilled and skilled households. That part of the calculation is fundamentally mechanical. The complicated part of the specification is the division of the population into unskilled and skilled labor portions and thus households. We have built formulations for that driven by the formal years of education attained by adults and the GDP per capita.

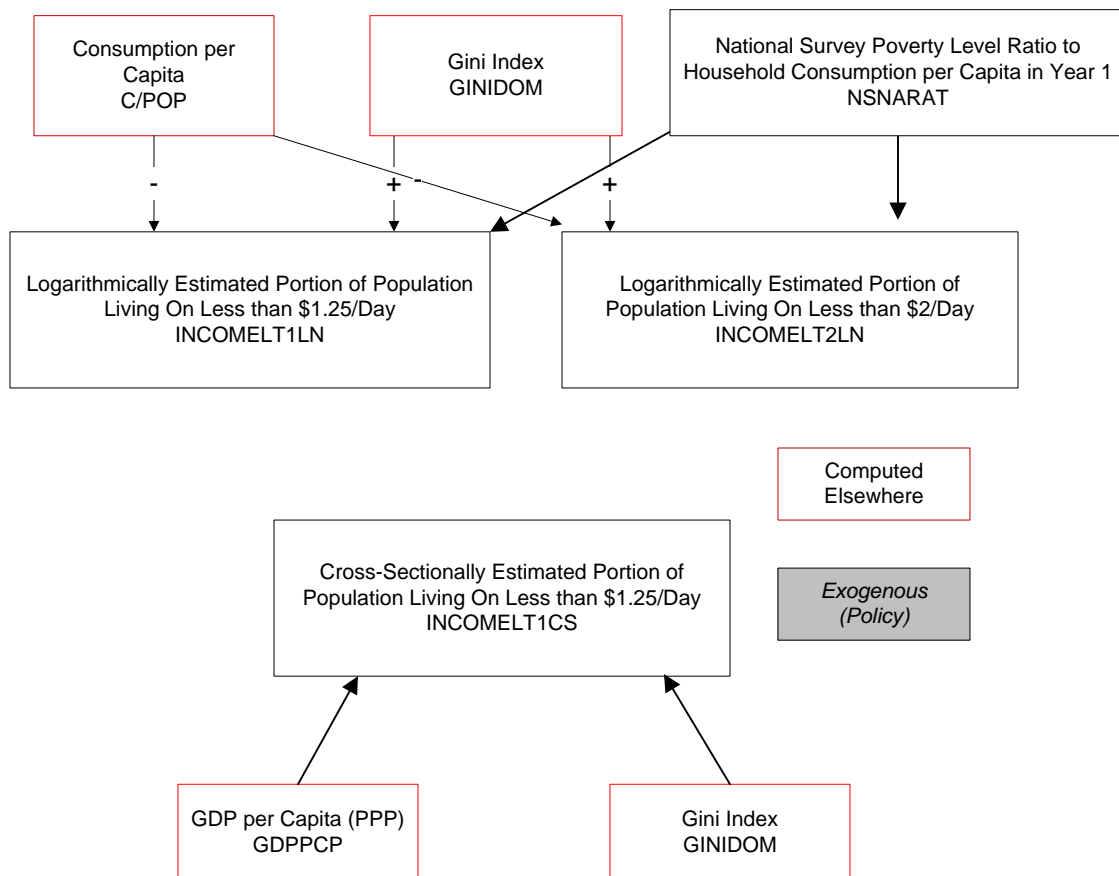


Given domestic Gini indices, it is also possible to compute global Gini indices, both treating countries as wholes (GINI) and computing across the world at the household level (GINIFULL).

2.12 Poverty

The calculation of poverty levels is fairly straightforward if one has the average level of consumption per capita (or income) and its distribution as indicated by the Gini index (and if one assumes that the distribution underlying the Gini index is log-normal). The internal calculation using those variables will, however, almost inevitably produce a rate of poverty at odds with the provided by national surveys. We therefore compute a ratio of those in the first year to allow adjustment in forecast years of the values from the lognormal calculation.

As a rough check on the values produced by lognormal calculation we also compute a value of poverty estimated from a cross-sectional function linking GDP per capita (and PPP) and Gini to rates of poverty.



2.13 Informal Economy

The IFs model represents the overall size of the informal economy in terms of the informal shares of total labor and total GDP. The model first uses several driver variables to calculate the informal labor share, broken down by sector (either the informal sector or the formal and household sector) and by sex. Years of educational attainment by adults (aged 15 and older) is used as the core or deep driver; the other drivers used represent the push and pull factors that keep or push people and businesses out of the formal sector or pull them into it. The push factors modelled are: (1) the business regulatory index, which represents the level of red tape facing businesses trying to formalize and the regulatory burden formal companies must operate under; (2) the tax rate on corporate profits; and (3) the level of corruption. The pull factors modelled are: (1) government spending on social welfare and pensions as a percent of GDP, which is used as a proxy for the general strength of government capacity and the social safety net which can keep workers from having to turn to informal means to survive; and (2) spending on research and development as a percent of GDP, which has an important and independent, though relatively weakly supported, effect on informal GDP share.

The informal labor share is then used, combined with additional drivers, to calculate the share of official GDP that is informal (we assume, based on the non-observed economy framework used by the OECD and a growing number of countries that official GDP statistics capture the size of the informal economy within it). The share of informal GDP has two direct forward linkages, (1) to multifactor productivity or total productivity, and (2) to tax revenues, with the implication that the greater the share of informal GDP, the slower official GDP grows² and the more constrained government revenues become. The impacts on GDP and government finances then feed forward to the rest of the IFs model, including back to the drivers of informal labor share.

² That is normally true, if the labor informal share exceeds the GDP informal share, implying that labor in the informal economy is less productive than that in the formal economy.

3. Equations

3.1 Overview

The growth portion of the goods and services module responds to endogenous labor supply growth (from the demographic model), endogenous capital stock growth (with a variety of influences on the investment level), and a mixture of endogenous and exogenous specification of advance in multifactor productivity (MFP). The endogenous portion of MFP represents a combination of convergence and country-specific elements that together create a conditional convergence formulation.

The equilibrium-seeking portion of the goods and services market module uses increases or decreases in inventories and prices (by sector) to balance demand and supply. Inventory stocks in each sector serve as buffers to reconcile demand and supply temporarily. Prices respond to stock levels. The central equilibrium problem that the module must address is maintaining domestic and global balance between supply and demand in each of the sectors of the model. IFs relies on three principal mechanisms to assure equilibrium in each sector: (1) price-driven changes in domestic demand; (2) price-driven changes in trade (IFs represents trade and global financial flows in pooled rather than bilateral form); and (3) stock-driven changes in investment by destination (changes in investment patterns could also be price-driven, but IFs uses stocks because of its recursive structure, so as to avoid a 2-year time delay in the response of investment).

The economic model makes no attempt through iteration or simultaneous solution to obtain exact equilibrium in any time point. In addition to being observationally obvious, Kornai (1971) and others have, of course, argued that real world economic systems are not in exact equilibrium at any time point, in spite of the convenience of such assumptions for much of economic analysis. Similarly, the SARUM global model (Systems Analysis Research Unit 1977) and GLOBUS (Hughes 1987) used buffer systems similar to that of IFs with the model "chasing" equilibrium over time.

Two "physical" or "commodity" models in IFs, agriculture and energy, have structures very similar to each other and to the economic model. They have partial equilibrium structures that optionally, and in the normal base case, replace the more simplified sectoral calculations of the goods and services market module.

The goods and services market sits within a larger social accounting matrix that tracks financial flows among households, firms, and the government. The integrated economic model also allows computation of income distribution and poverty rates.

3.2 The Goods and Services Market

3.2.1 *The Production Function: Basic Overview*

Cobb-Douglas production functions involving sector-specific technology or multifactor productivity (TEFF), capital (KS) and labor (LABS) provide potential value added (VADDP) in each sector, taking into account the level of capacity utilization (CAPUT), initially set exogenously (*caputtar*). In a multi-sector model the functions require sectoral

exponents for capital (CDALFS) and labor that, assuming constant returns to scale, sum to one within sectors.

Solow (1956) long ago recognized that the standard Cobb-Douglas production function with a constant scaling coefficient in front of the capital and labor terms was inadequate because the expansion of capital stock and labor supply leave a large portion of economic growth unexplained. It then became standard practice to represent an exogenously specified growth of technology term in front of the capital and labor terms as "disembodied" technological progress (Allen 1968: Chapter 13). Romer (1994) began to show the value of unpacking such a term and specifying its elements in the model, thereby endogenously representing this otherwise very large residual, which we can understand to represent the growth of multifactor productivity (MFP).

In IFs that total endogenous productivity growth factor (TEFF) is the accumulation over time (hence a stock like labor and capital) of annual values of growth in multifactor productivity (MFPGRO). There are many components contributing to growth of productivity, and there is a vast literature around them. See, for example, Barro and Sala-i-Martin (1999) for theoretical and empirical treatment of productivity drivers; also see Barro (1997) for empirical analysis (or McMahon 1999) for a focus on education.

In the development of IFs there was a fundamental philosophic choice to make. One option was to keep the multi-factor productivity function very simple, perhaps to restrict it to one or two key drivers, and to estimate the function as carefully as possible. Suggestions included focusing on the availability/price of energy and the growth in electronic networking and the knowledge society.

The second option was to develop a function that included many more factors known or strongly suspected to influence productivity and to attempt a more stylistic representation of the function, using empirical research to aid the effort as much as possible. The advantages of the second approach include creating a model that is much more responsive to a wide range of policy levers over the long term. The disadvantages include some inevitable complications with respect to overlap and redundancy of factor representation, as well as some considerable complexity of presentation.

Because IFs is a thinking tool and an extensively integrated multi-model system, the second approach was adopted, and the production function has become an element of the economic model that will be subject to regular revision and enhancement. IFs groups the many drivers of multifactor productivity into five categories, recognizing that even the categories overlap somewhat. The base category is one that represents core technological development and transfer elements of convergence theory, with less developed countries gradually catching up with more developed ones. The four other categories incorporate factors that can either retard or accelerate such convergence, transforming the overall formulation into one of conditional convergence.

The convergence base. The base rate of multifactor productivity growth (MFPRATE) is the sum of the growth rate for technological advance or knowledge creation of a technological leader in the global system (*mfpleadr*) and a convergence premium

(MFPPrem) that is specific to each country/region. The basic concept is that it can be easier for less developed countries to adopt existing technology than for leading countries to develop it (assuming some basic threshold of development has been crossed). The base rate for the leader remains an unexplained residual in the otherwise endogenous representation of MFP, but that has the value of making it available to model users to represent, if desired, technological cycles over time (e.g. Kondratief waves). The base also includes a correction term (MFPCor) that is initially set to the difference between empirical growth of MFP (calculated the first year as a residual between factor growth and output growth) and the sum of the technological leader and convergence premium terms. Over time, the correction term is phased out, but the four other terms (below) become key drivers of country-specific productivity. In fact, significant change in the other terms can either undercut the foundational convergence process or greatly augment it.

Human capital. This term has multiple components, including changes in educational spending as a portion of GDP, educational attainment of adults, and changes in health expenditure. For example, Barro and Sala-i-Martin (1999: 433) estimated that a 1.5% increase in government expenditures on education translates into approximately a 0.3% increase in annual economic growth.

Social capital. Similarly, social capital representation aggregates several components including economic freedom and absence of overt social conflict. Illustratively, the value of the parameter for economic freedom (*elgref*) was estimated in a cross-sectional relationship of change in GDP level from 1985 to 1995 with the level of economic freedom. Similarly, Barro places great emphasis in his estimation work on the "rule of law".

Physical capital. In collaborative work with the IFs project, Robert Ayres correctly emphasized the close relationship between energy supply availability and economic growth. For instance, a rapid increase in world energy prices (WEP) essentially makes much capital stock less valuable. IFs uses world energy price relative to world energy prices in the previous year to compute an energy price term. The physical capital term also represents the extent of various types of infrastructure in a society.

Knowledge Capital. This fourth term includes changes in the R&D spending, computed from government spending (GDS) on R&D as a portion of total government spending (GOVCON) contribute to knowledge creation, notably in the more developed countries. Gliberman (2000) reviewed empirical work on the private and social returns to R&D spending and found them to be in the 30-40% range; see also Griffith, Redding, and Van Reenen (2000). Many other factors undoubtedly contribute to superior knowledge development diffusion. This term represents especially the extent of economic integration with the global economy via trade.

All of the elements computed in the human, social, physical and knowledge capital terms are used in shaping economic productivity and growth rates of the model on a differential basis—that is, they are computed and evaluated relative to underlying "expected" patterns given overall economic development levels. Their actual levels can be above or below

expected ones and they can therefore either add to or slow down the productivity growth rate. See the detailed equations of the production function for elaboration.

3.2.2 *The Production Function: Detail*

The core equation for the production function computes value added (VADD) from technology (TEFF), capital (KS), labor (LABS), and capacity utilization (CAPUT), with a time-constant scaling parameter (CDA) assuring that gross production is consistent with data the first year. GDP is, of course, the sum across sectors of valued added. Although the production function can serve all sectors of IFs, the parameters *agon* and *enon* act as switches; when their values are one, production in the agricultural and primary energy sectors, respectively, are determined in the larger, partial equilibrium models and the values then override this computation (see documentation on those models for detail).

$$VADDP_{r,s} = CDA_{r,s,t=1} * TEFF_{r,s} * CAPUT_{r,s} * KS_{r,s}^{CDALFS_{r,s}} * LABS_{r,s}^{(1-CDALFS_{r,s})}$$

$s = 3,4 \dots nsectr$ or $s = 1,2 \dots nsectr$

where

$$CDA_{r,s,t=1} = \frac{VADDP_{r,s,t=1}}{(KS_{r,s,t=1})^{CDALFS_{r,s,t=1}} * (LABS_{r,s,t=1})^{(1-CDALFS_{r,s,t=1})}}$$

$$CAPUT_{r,s,t=1} = \mathit{caputtar}$$

Other topics in our documentation explain the dynamics underlying change in capital stock (through investment and depreciation) and labor supply (through demographic change and participation patterns). The rest of this topic will focus on the computation of the elements that go into the MFP or technology term (TEFF), working our way progressively deeper into their determinants.

At the most basic level the stock term is simply an accumulation of the annual increments in MFP.

$$TEFF_{r,s} = TEFF_{r,s,t=1} * (1 + MFPGRO_{r,s}) \text{ where } TEFF_{r,s,t=1} = 1.0$$

The annual growth in MFP (MFPGRO) consists of a base rate linked to systemic technology advance and convergence (MFPRATE) plus four terms that affect MFP growth over time as a result of human, social, physical, and knowledge capital.

$$MFPGRO_{r,s} = MFPRATE_{r,s} + MPFHC_{r,s} + MFPPSC_{r,s} + MFPPC_{r,s} + MFPKN$$

Each of the following subsections elaborates one of the five terms.

Annual Base Technology Growth in MFP

The base rate related to technology or other factors unexplained by the four capital terms is anchored by an exogenous specification of the sector-specific rate of advance in the

systemic leader's technology (*mfpleadr*); the leader is assumed to be the United States. (The rate in the leader's ICT sector gradually converges over time to that of the service sector.) On top of that is an endogenously computed premium computed for convergence of each country/region (MFPPrem), a term that is a function of GDP per capita at PPP, the function for which posits an inverted-V shape with the greatest potential for technological convergence to the leader among middle-income countries. In this calculation there needs to be a correction factor (MFPCor), computed to assure that the actual rate of this growth is consistent with the actual amount of growth for a country/region (MFPGRO) that is unexplained by capital and labor growth in the first model run year. That correction factor converges to 0 over the number of years specified by *mfpcnv*. This convergence assumption has significant implications for model behavior because it tends to slow down growth in countries (like China) that have had a burst of growth beyond that which the rates of the leader and the catch-up factor would lead us to anticipate and to speed up growth in countries (like the transition states of Central Europe) that have suffered a reduction similarly unexpected by the basic formulation.

$$MFPRATE_{r,s} = mfpleadr_s + MFPPrem_r + MFPCor_{r,s} + MFPGloCor_{t=1} + mfpcnv + mfpcnv * zy + mfpcnv$$

where

$$MFPPrem_r = Func(GDPPCP_r)$$

$$MFPCor_{r,s,t=1} = MFPGRO_{r,s,t=1} - MFPRATE_{r,s,t=1} \text{ and}$$

$$MFPCor_{r,s,t} = (ConvergeOverTime(MFPCor_{r,s,t=1}, 0, mfpcnv))$$

$$MFPGloCor_{t=1} = \frac{\sum_{r=1}^R \sum_{s=1}^S MFPCor_{r,s,t=1} * VADD_{r,s,t=1}}{\sum_{r=1}^R \sum_{s=1}^S VADD_{r,s,t=1}}$$

It is possible for the regionally and sectorally specific MFP correction terms (MFPCOR) to still leave a small discrepancy between the global economic growth in the data and the initial growth rate in the model. To avoid this, we also compute a global correction factor (MFPGloCor) as the region/country and value added weighted sum of the individual country-sector terms divided by the global and regional sums of value added.

In addition to the basic technology terms and the correction factors, there are three parameters in the MFPRATE equation (with zero values as the default) that allow the model user much control over assumptions of technological advance. The first is basic parameter (*mfpcnv*) that allows a global growth increment or decrement; the second is a parameter (*mfpcnv*) that allows either a constant rise or slowing of growth rate globally, year by year, where *zy* is the count of the model run years across time; finally is a frequently-used parameter (*mfpcnv*) allowing flexible intervention for any country/region.

Driver Cluster 1: Human Capital

The general logic of each the four driver clusters around human, social, physical, and knowledge capital is the same. Each cluster aggregates several variables that generally contribute to productivity. For each variable, such as average years of adult education in the human capital cluster, there is an expected value and an actual value. It is the difference between actual and expected values that gives rise to a positive or negative contribution to productivity and growth. Most expected values are identified in a relationship with GDP per capita at PPP.

That is, there is a tendency for most developmentally supportive variables to advance in a rough relationship with each other and with GDP per capita (Kuznets 1959 and 1966; Chenery and Syrquin 1975; Syrquin and Chenery 1989; Sachs 2005). To the degree that they do, such advance can be understood to be consistent also with the overall technological advance of the country. If, however, a variables such as years of formal education attained by adults exceeds the typical or expected value for a country at a given level of GDP per capita, we can expect that variable to add something more to productivity. Similarly, falling behind the expected value could retard productivity advance. To illustrate and emphasize this point, even a country for which adult education levels advance could find that education is not keeping up with the advance in other developmental variables including GDP per capita and find that its education levels move from contributing to productivity enhancement to decrementing that productivity enhancement.

In the human capital cluster there are six variables that add to or subtract from the human capital (MFPHC) term: the educational spending contribution (EdExpContrib), the years of adult education contribution (EdYearsContrib), the boost from life expectancy years (LifExpEdYrsBoost) assumed to generate (via *mfpedlifexp*) extra years of education, the stunting contribution (StuntContrib) related to undernutrition of children, the disability contribution (DisabContrib) related to morbidity from the health model, and vocational education contribution (edVoccontrib) resulting from growth (or decline) in vocational share of lower and upper secondary enrollment,. The first five of these six drivers have a similar formulation while the formulation for vocational education is slightly different. The first five, as computed in IFs often as a result of extended formulations and even other models of system (as with life expectancy, computed in the health model) are compared with an expected value. In the case of disability, the expected value is set to the world average level (WorldDisavg), but all other expected values (EdExpComp, EdYrsComp, LifExpComp, and StuntingComp) are computed as functions of GDP per capita at PPP. As the provision of vocational education does not follow any common pattern or trend and is rather a matter of policy choices made (or will be made) by the particular country, it was not possible to calculate an expected value for this variable We have instead computed the vocational education contribution from changes in vocational share over time with appropriate moving averages to capture the lag required in materializing such contribution and to smooth out the contribution over time. (Note: In the base case of the model, vocational shares do not change and as such EdVocContrib is zero).

In each case a parameter drawn from study of the literature and/or our own analysis converts the difference between actual and expected into a positive or negative contribution to MFP. (Because of the recursive structure of IFs, some terms rely on variables from the previous time step, estimated from the current time step with a correction factor based on initial GDP growth.)

$$MFPHC_r = EdExpContrib_r + EdYrsContrib_r + LifExpEdYrsBoost_r + StuntContrib_r + DisabContrib_r + edVocContrib_r$$

where

$$EdExpContrib_r = \left(\frac{GDS_{r,s=EDUC,t-1} * (1 + IGDPRCor_r)}{GDP_{r,t-1} * (1 + IGDPRCor_r)} * 100 - EdExpComp_r \right) * \mathbf{mfpedspn}$$

$$EdYrsContrib_r = (EDYRSAG15_{r,t-1} - EdYrsComp_r) * \mathbf{mfpedyrs}$$

$$LifExpEdYrsBoost_r = (Amin(LIFEXP_{r,t-1}, SLifExpSlow_r + 2) - LifExpComp_r) * \mathbf{mfpedlifexp} * \mathbf{mfpedyrs}$$

$$StuntContrib_r = (HLSTUNT_{r,t-1} - StuntingComp_r) * \mathbf{mfpstunt}$$

$$DisabContrib_r = \left(Amin \left(0.17, \frac{SumHIYLDWork_r}{Amax(0.001, POPWORKING_r)} \right) - WorldDisAvg \right) * \mathbf{mfphlyld}$$

$$edVocContrib_{r,t} = f(EDSECLOWRVOC_{r,t-1}, EDSECLOWRVOC_{r,t-2}, EDSECUPPRVOC_{r,t-1}, EDSECUPPRVOC_{r,t-2})$$

$$edVocContrib_r = (0.9 * edVocContrib_{r,t-1} + 0.1 * edVocContrib_{r,t}) * \mathbf{mfpedvoc}$$

where

Vocational contribution for lower secondary is calculated as:

$$eSecLowrVoc_{r,t} = 0.9 * eSecLowrVoc_{r,t-1} + 0.1 * EDSECLOWRVOC_{r,total,t-1}$$

$$edVocSecLowrContrib_r = EDSECLOWRVOC_{r,total,t-1} - eSecLowrVoc_r$$

A similar calculation is done for upper secondary vocational and the two are averaged as shown below:

$$edVocContrib_r = (edVocSecLowrContrib_r + edVocSecLowrContrib_r)/2$$

Driver Cluster 2: Social Capital

The logic of comparison of actual with expected values is the same as that described above for human capital in the case of the six factors that contribute to social capital: economic freedom as in the Fraser Institute measure (EconFreeContrib), government effectiveness as in the World Bank measure (GovtEffContrib), corruption as in the Transparency International measure (CorruptContrib), democracy as in the Polity project measure (DemocPolityContrib), freedom as in the Freedom House measure (FreedomContrib), and conflict as in the IFs project's own measure tied in turn to the work of the Political Instability Task Force (ConflictContrib). In each case other than that for conflict, the expected values (EconFreeComp, GovEffectComp, CorruptComp, DemocPolityComp, and FreeComp) are computed from functions with GDP per capita at PPP. In the case of conflict, the expected value is set at the initial year's value (and the comparison is reversed because lower conflict values contribute positively to MFP).

$$MFPSC_r = EconFreeContrib_r + GovtEffContrib_r + CorruptComp_r + DemocPolityContrib_r + FreedomContrib_r + ConflictContrib_r$$

where

$$EconFreeContrib_r = (ECONFREE_{r,t-1} - EconFreeComp_r) * mfpfree$$

$$GovtEffectContrib_r = (GovEffect_{r,t-1} - GovEffectComp_r) * mfpgoveff$$

$$CorruptContrib_r = (GOVCORRUPT_{r,t-1} - CorruptComp_r) * mfpgovcor$$

$$DemocPolityContrib_r = (DemoPolity_{r,t-1} - DemcPolityComp_r) * mfpdemoc$$

$$FreedomContrib_r = (FREEDOM_{r,t-1} - FreeComp_r) * mfpfree$$

$$ConflictContrib_r = (GOVINDSECRI_{r,t=1} - GOVINDSECR_{r,t}) * mfpconflict$$

Driver Cluster 3: Physical Capital

The logic of the physical capital cluster is again parallel to that of the human and social capital clusters and involves the comparison of an actual (that is, IFs computed) with an expected value. The formulation for MFPPC can actually take several forms depending on the value of a switching parameter (*inframfpsw*) but the standard form involves four contributions, from traditional infrastructure (InfraTradContrib), ICT infrastructure (InfraICTContrib), other infrastructure spending level (InfOthSpndContrib), and the price of energy (EnPriceTerm). The last term is included because higher prices of energy can make some forms of capital plant no longer efficient or productive.

In the case of this cluster only the expected value of the traditional infrastructure index (InfraIndTradComp) and the expected value of other infrastructure spending (InfraOthSpndComp) are computed as most other cluster elements are, namely as a function of GDP per capita at PPP. In the case of the ICT index contribution, the technology has been evolving so rapidly that there is not really a basis for an expected value with some stability over time. Instead the contribution from ICT is computed in terms of a moving average value of change over time, such that faster rates of change contribute more to MFP as the moving average expected value lags further behind the actual. In the case of the energy price term, the "expected" value is set equal to the energy price in the first year of the model run. As with other clusters and the variables in them, a single parameter links the discrepancy between actual and expected values to MFP.

$$MFPPC_r = InfraTradContrib_r + InfraICTContrib_r + InfOthSpndContrib_r + EnPriceTerm_{r,t-1}$$

where

$$\begin{aligned} InfraTradContrib_r &= (INFRAINDTRAD_{r,t-1} - InfraIndTradComp_r) \\ &* \mathbf{mfpinfraindtrad} \end{aligned}$$

$$\begin{aligned} InfraICTContrib_r &= (0.8 * IndICTIndChange_{r,t-1} + 0.2 * IndICTIndChange_{r,t}) \\ &* \mathbf{mfpinfraindict} \end{aligned}$$

where

$$IndICTIndChange_r = InfraIndICT_{r,t-1} - InfraIndICT_{r,t-2}$$

$$\begin{aligned} InfraOthSpndContrib_r &= \left(\frac{GDS_{r,s=InfraOther,t-1} * (1 + IGDPRCor_r)}{GDP_{r,t-1} * (1 + IGDPRCor_r)} * 100 \right. \\ &\left. - INFRAOthSpndComp_r \right) * \mathbf{mfpinfrothspnd} \end{aligned}$$

$$PhysicalCapitalTerm_{r,s} = EnPriceTerm_{t-1}$$

$$EnPriceTerm_{t-1} = \left(\frac{WEP_{t-1} - WEP_{t=1}}{WEP_{t=1}} \right) * mfpnpri$$

Driver Cluster 4: Knowledge Capital

Following the pattern of other MFP driver clusters, the one for knowledge accumulation includes terms that compare actual model and typical or expected values and use parameters to translate the differences into increments or decrements of MFP. In this case the three terms represent R&D spending (RDExpContrib), economic integration via trade with the rest of the world (EIntContrib) and the share of science and engineering among all tertiary degrees earned (EdTerSciContrib). In the first and third case the expected value (RDExpComp; EdTerGRSciEnComp) is a function of GDP per capita at PPP. In the second instance, there is no clear relationship between extent of economic integration and GDP per capita, so the model compares a moving average of trade openness (exports plus imports as a percentage of GDP) with the initial value of openness (because trade is computed later in the computational sequence for each year, the values of trade variables lag one year behind those of the production function). Given the extreme global range of trade openness, the elasticity term itself in this relationship is variable, with values decreasing when initial openness is greater (that is, countries that start with less openness gain more from the same percentage point increases in it).

$$KnowledgeTerm_{r,s} = RDExpContrib_r + EIntContrib_r + EdTerSciContrib_r$$

where

$$RDExpContrib_r = (RANDEXP_{r,t-1} - RDExpComp) * mfprandd$$

$$EIntContrib_r = (TradeTermMA_{r,t-1} - TradeTerm_{r,t=1})/10 * ElastTerm_r$$

where

$$TradeTerm_r = \frac{X_r + M_r}{GDP_r} * 100$$

$$TradeTermMA_r = 0.8 * TradeTermMA_{r,t-1} + 0.2 * TradeTerm_{r,t-1}$$

$$ElastTerm_r = Amax \left(0.1 * mfpeconint, Amin \left(2 * mfpeconint, mfpeconint * \left(1 + \frac{50 - TradeTerm_{r,t=1}}{50} \right) \right) \right)$$

$$\begin{aligned}
EdTerSciContrib_r &= (EDTERGRSCIEN_{r,t-1} * EDTERGRATE_{r,t-1}/100 \\
&- EdTerGrSciEnComp) * mfpedscien
\end{aligned}$$

Issues Concerning Parameterization and Interaction Effects

Although our approach to calculation of MFP creatively connects developments in many other models in the IFs system to it, parameterization of the effects individually and in interaction is complicated and uncertain. Hughes (2005) documented the original creation of the structure and its parameterization based on existing literature.

One of the concerns with this approach is the possibility of double counting of effects from the large number of variables fed into the MFP calculations. In general the project has dealt in part with this by selecting conservative values for the parameters when studies indicate possible ranges of contribution of the variables to productivity and/or growth. Another concern is that a very large or extreme advance by one or a small subset of variables could have inappropriately large impacts of productivity given the fundamental conceptual foundation of the approach in the notion that development involves widespread and reinforcing structural changes across many variables. In order to limit this possibility, we have created an algorithmic function (MFPContribAdj) to adjust the multiple MFP contributions and dampen especially high positive or negative contributions of the four cluster terms (MFPHC, MFPS, MFPPC, and MFPKN).

The Relationship of Physical Models to the Economic Model

IFs normally does not use the economic model's equations representing MFP and production for the first two economic sectors, because the agriculture and energy models provide gross production for them (unless those sectors are disconnected from economics using the *agon* and/or *enon* parameters). Instead, the two physical models provide gross production, translated to value terms, back to the economic model. See Section 3.2.3 for discussion of gross production.

In addition to this impact of the physical models on the economic model, there is one more of importance. Physical shortages on energy may constrain actual value added in each sector (VADD) relative to potential production. Economists typically do not accept such shortages as a real world phenomenon because (at least in theory) prices rise to clear markets; yet periods like the 1970s when governments intervened in those markets, such shortages do appear and they can in some IFs scenarios. In those situations, IFs assumes that energy shortages (ENSHO), as a portion of domestic energy demand (ENDEM) and export commitments (ENX) lower actual production in all sectors through a physical shortage multiplier factor (SHOMF). A parameter/switch (*squeeze*) controls this linkage and can turn it off.

$$VADD_{r,s} = VADDP_{r,s} * (1 - ShoMF_r) * \left(\frac{MKAV_{s=manuf}}{MKAV_{s=manuf,t=1}} \right)^{prodme}$$

where

$$ShoMF = \frac{ENSHO_r}{ENDEM_r + ENX_r} * squeeze$$

In addition, the translation of potential into actual production depends on the imports of manufactured goods (MKAV), which serve as a proxy for both availability of intermediate goods and for technological imports. A parameter (PRODME) also controls this relationship.

Although the IFs model represents prices in real terms (no monetary sector and no inflation), there are relative sectoral price changes (PRI). Some of those can be quite dramatic over time, especially in the agricultural and energy sectors where the equilibration of physical representations of supply and demand can swing those prices. Such relative price swings can, in the real world, give added drag or boost to the value added in certain sectors and for economies as a whole. To represent this we compute a relative-price adjusted version of value added (VADDRPA), which is the normal value added weighted by world sectoral prices (WP); the prices are lagged a year because of the recursive model structure.

$$VADDRPA_{r,s} = VADD_{r,s} * WP_{s,t-1}$$

3.2.3 Gross Production and Intersectoral Flows

Given value added in each sector it is possible using an exogenously provided input/output matrix (A) to determine the level of gross production in each sector. We will see below, however, that the A-matrix is actually computed as a function of GDP per capita—the model interpolates among multiple A-matrices for various levels of GDP per capita in a procedure originally developed for the GLOBUS model (Hughes 1987).

$$ZS_{r,s} = \frac{VADD_{r,s}}{1 - \sum^{Row} A_{row,column=s}}$$

Given gross production and the A-matrix we can compute intersectoral flows (INTS).

$$INTS_{row,column=s} = ZS_{r,s} * A_{row,column=s}$$

Production available for final demand (PFD) is the residual of gross production minus the sum delivered to all columns.

$$PFD_{r,s} = ZS_{r,s} - \sum^{Column} INTS_{row=s,column}$$

GDP per capita-linked generic A-matrices are created in the pre-processor of IFs. To build them the IFs project turned to the IO matrices collected in the Global Trade Analysis Project. That database includes extensive data, including IO matrices, for 127 regions/individual countries across 57 sectors in GTAP 8 (every version tends to increase

geographic coverage). With origins in 1992 of the now global project at the agricultural economics department of Purdue, GTAP heavily represents agricultural sectors. Narayanan, Aguilar and McDougall (2012) documented GTAP 8 (see also earlier versions including Dimaranan and McDougall (2002) for GTAP 5).

The pre-processor of IFs (Hughes and Irfan 2006) processes the IO matrices of GTAP to create those needed by IFs. One aspect of that involves collapsing the sectors to those of IFs (currently 6) using a concordance table. Even more importantly, however, the pre-processor generates from the most recent GTAP files within IFs a set of nine generic IO matrices to represent the average technical coefficient pattern of countries at different levels of GDP per capita. The generic matrices are calculated as unweighted averages of matrices for all countries with GDPs per capita in categories established by lower-end breakpoints of \$0, \$175, \$375, \$750, \$1,500, \$3,000, \$6,000, \$12,000, and \$24,000.

The assumption behind the generic IO matrices is that countries at different GDP per capita levels typically use different types of technology. The resultant IO matrices bear this out in ways that seem intuitively plausible. For example, Tables 3.1 and 3.2 show (using earlier data from GTAP 5) the technical coefficient matrices for extreme levels of GDP/capita, below \$100 and above \$24,000, respectively. Note, for instance, how much lower a share of manufactures goes into the agricultural sector in the richest countries relative to the poorest, and how much more of the IC sector goes back into the IC sector in richer countries.

	AG	RM	PE	MN	SR	IC
AG Sector	0.2624	0.0112	0.0008	0.0846	0.0194	0.0014
RM Sector	0.0041	0.0425	0.1571	0.0499	0.0087	0.0418
PE Sector	0.0048	0.2158	0.0265	0.0735	0.0119	0.0362
MN Sector	0.0522	0.0540	0.0687	0.1652	0.0774	0.0780
SR Sector	0.1847	0.2260	0.2177	0.1797	0.1721	0.1808
IC Sector	0.0026	0.0090	0.0040	0.0058	0.0105	0.0271

Table 3.2 Generic IO Matrix for Countries with GDP/Capita Above \$24,000

	AG	RM	PE	MN	SR	IC
AG Sector	0.3483	0.0005	0.0017	0.0133	0.0107	0.0004
RM Sector	0.0132	0.0366	0.1385	0.0186	0.0063	0.0067
PE Sector	0.0141	0.2660	0.0118	0.0823	0.0040	0.0287
MN Sector	0.0645	0.0614	0.0822	0.1812	0.0670	0.0856
SR Sector	0.1586	0.1786	0.1533	0.2004	0.2399	0.1632
IC Sector	0.0061	0.0093	0.0113	0.0156	0.0169	0.0966

These generic matrices are used for two purposes. First, they are used for estimating values for countries of IFs that are NOT in the GTAP data set. Second, they are used in the actual dynamic calculations of the model. As countries rise in GDP/capita, interpolations between matrices above and below their level allow us to gradually change the matrix representing each country.

GTAP also provides data on return to four factors of production in each sector: land, unskilled labor, skilled labor, and capital. These returns represent value added and are very important data for the value added blocks of the SAM. The pre-processor also collapses these values into the (six) sectors of IFs and computes generic shares of the factors in value added by GDP per capita category, using the same unweighted average technique used for the IO coefficients. Once again the generic value-added shares are used both to fill country holes in the GTAP data set and to provide a basis for dynamically representing changes in those shares as countries develop.

Table 3.3 Generic Returns to Labor for Countries Below GDP/Capita \$100

	AG	RM	PE	MN	SR	IC
Unskilled	0.2873	0.1544	0.1763	0.1378	0.2878	0.1721
Skilled	0.0063	0.0264	0.0384	0.0234	0.1325	0.1132

Table 3.4 Generic Returns to Labor for Countries Above GDP/Capita \$24,000

	AG	RM	PE	MN	SR	IC
Unskilled	0.1795	0.2146	0.0854	0.2159	0.2074	0.2054
Skilled	0.0386	0.0842	0.0952	0.1060	0.1715	0.1618

The changes across the levels of GDP/capita appear reasonable. Note, for instance, the general shift of return to skilled from unskilled labor and the increase in returns to labor in total for the manufacturing and ICT sectors (at the expense of capital and other inputs).

The equations for value added and income in household categories can be found in the discussion of households as agents. Although the GTAP data by no means provide everything that was needed for the generation of universal SAMs, the project is aware of the utility of SAMs (Brockmeier and Arndt 2002) and provided several primary data inputs that serve our purposes.

Further, the GTAP data and the GDP per capita threshold approach is used in the pre-processor to compute a generic table for labor demand coefficients by sector, distinguishing skilled and unskilled labor. These are stored in the IOLaborCoefs table of IFs.mdb and are available for the computation of labor demand in the economic model.

3.2.4 Labor Supply

Labor supply (LAB) is a function of population, depending on a labor force participation rate (LAPOPR). In earlier versions of IFs that participation rate was an exogenous parameter. It has now been decomposed into three elements: (1) the share of the population in the traditional working ages between 15 and 65, (2) the retirement age, and (3) the participation rate of women.

At one time we computed the share of population in the traditional working ages as a ratio of the size of the population in that age category (POP15TO65) and population (POP).

$$Pop15to65Ratio_r = \frac{POP15TO65_r}{POP_r}$$

Although the name of the variable has not been changed yet, the formulation has now been changed in order to represent variable ages of entry into the work force (*workageentry*) and retirement age (*workageretire*) across countries and time. Those parameters are used to compute an actual working aged population (POPWORKING) or potential labor force in an algorithm across population age categories in the population model.

$$Pop15to65Ratio_r = \frac{POPWORKING_r}{POP_r}$$

where

$$POPWORKING_r = F(AgeDst_{c,r}, \mathbf{workageentry}_r, \mathbf{workageretire}_r)$$

The participation of women in the work force (FEMSHRLAB) as a share of the total labor force is assumed to grow over time with an exogenous parameter based on past experience (**femshrgr**); the representation of women as a share of the labor force rather than in terms of percentage points of their participation somewhat complicates the equations below. The growth in labor-force share is modified by a multiplier (FemShrLabMul) that introduces saturation as female participation rates approach a target (FemShrTar). The model user can modify that target, normally between 50 and 60% via an exogenous parameter (**labfemshrm**).

$$FEMSHRLAB_r = FEMSHRLAB_{r,t=1} + \mathbf{femshrgr} * FemShrLabMul_r$$

where

$$FemShrLabMul_r = \frac{FemShrTar_r - FEMSHRLAB_{r,t=1}}{FemShrTar_r - FEMSHRLAB_{r,t=1}}$$

$$FemShrTar_r = \text{Amin}(60, \text{Amin}(50, FEMSHRLAB_{r,t=1} + 20) * \mathbf{labfemshrm}_e)$$

Given the three drivers of labor force participation rates (LAPOPR) it is possible to compute it relative to the rate in the first year:

$$LAPOPR_r = \left(LAPOPR_{r,t=1} + \frac{FEMSHRLAB_r - FEMSHRLAB_{r,t=1}}{100} * \frac{FEMSHRLAB_{r,t=1}}{100} \right) * \frac{Pop15to65Ratio_r}{Pop15to65Ratio_{r,t=1}} * \mathbf{labretagem}_r * \mathbf{lapoprm}_r$$

Change in the retirement age can be reflected in the above equation via a multiplicative parameter (**labretagem**), but it does not give the precision of control that **workageretire** does and is no longer recommended for use. The multiplier on participation rate (**lapoprm**) is, however, a useful scenario intervention point.

The product of participation rate and population provides the total labor pool (LAB).

$$LAB_r = POP_{r,t} * LAPOPR_r$$

The total labor pool can be divided into subcomponents in two different ways. First, labor is spread across production sectors (LABS). Second, it is differentiated by household type in to skilled and unskilled labor (LABSUP).

Labor Supply by Sector

Labor by sector of the economy (LABS) is a share of the total labor force (LAB) minus unemployment calculated at an exogenous unemployment rate (UNEMPR). The sectoral share is calculated in a function that estimates the labor demand for each unit of value added (VADD) at given levels of GDP per capita (GDPPC).

$$LABS_{r,s} = \frac{LAB_r * (1 - unempr_r) * \sum^H LaborF_{r,s}}{\sum^S \sum^H LaborF_{r,s}}$$

where

$$LaborF_{r,s} = AnalFunc(GDPPC_r, VADD_{r,s})$$

Labor Supply by Household Skill Category

Labor by household type (LABSUP) is determined by calculation of the percentage that comes from skilled households (PerSkilled). At the core of that calculation of percent skilled is an analytical function with GDP per capita at PPP (GDPPCP). But the analytical function is fundamentally tied to conditions that prevail in the contemporary era (as captured by the GTAP data on which it is based). We know that education levels have been increasing, even to the extent of involving some credential inflation. Thus the percentage of the labor force considered skilled is likely to increase faster than the function indicates. Thus it is modified by a skilled labor adjustment factor (LabSupSkiAdj) that takes into account the difference between the actual extent of education among adults over 15 (EDYEARS15) and the expected level and translated that into a boost or reduction in the percentage skilled in the same manner that differences between actual and expected years affect MFP in the production function. The YearsEdDiff term simply maintains the difference by country/region of the initial (year 2) difference between the expected years from the function and the actual years of education.

$$LABSUP_{r,h=skilled} = LAB_r * \frac{PerSkilled}{100}$$

where

$$PerSkilled = AnalFunc(GDPPCP_r) * (1 + LabSupSkiAdj)$$

$$LabSupSkiAdj = \frac{EDYEARS15_r - YearsEdExp}{YearsEd} * 0.3$$

and

$$YearsEd = AnalFunc(GDPPCP_r)$$

$$YearsEdDiff_r = EDYRSAG15_{r,t=2} - YearsEd_{t=2}$$

$$YearsEdExp = YearsEd + YearsEdDiff_r$$

The unskilled labor is computed as a residual.

$$LABSUP_{r,h=Unskilled} = LAB_r - LABSUP_{r,h=Skilled}$$

3.2.5 Capital-Labor Shares

The Cobb-Douglas exponent (CDALF) of the production function is known to change over long time periods, giving somewhat less weight to capital as an economy becomes more capital intensive (see, for instance, Thirlwall 1977: chapter 2). The GTAP project data again provided the basis for an estimation of this relationship.

That function, in combination with data from Thirlwall and others on sectoral differences in capital share, allow the computation of sectoral capital shares (CDAlfS), normalized so as to generate CDALF for the total economy. A parameter (*salpha*) represents a generic pattern of capital share variation across production sectors.

$$CDALF_r = AnalFunc(GDPPCP_r)$$

$$CDAlfS_{r,s} = Normalization(CDALF_r, salpha_s)$$

3.2.6 GDP at MER and PPP and Relative-price Adjusted GDP

Gross regional or domestic product (GDP) is the sum of value added across sectors, which would also equal the sum of production for final demand across sectors.

$$GDP_r = \sum^s VADD_{r,s}$$

The GDP per capita (GDPPC) and the regional economic growth rate (GDPR) follow easily.

$$GDPPC_r = \frac{GDP_r}{POP_r}$$

$$GDPR_r = \frac{GDP_r - GDP_{r,t-1}}{GDP_{r,t-1}}$$

The basic GDP figures for the model are represented in dollars at official exchange rate values. It is important, however, to estimate the value of GDP and GDPPC at purchasing power parity levels as well (GDPP and GDPPCP). To do that we need to compute a purchasing power parity conversion value (PPPConV). Data sources provide the initial conversion value. IFs uses an analytic function that relates GDP per capita at MER to that at PPP to compute change in the conversion value over time.

The initial condition for the conversion ratio comes from data values for the GDP per capita at PPP and MER. Over time, the conversion ratio (which can be considerably

above 1 for developing countries) erodes. The annual erosion is the difference between the previous year's value and a product of two terms: (1) an initial ratio that controls for any difference between the actual GDP per capita at PPP and the value produced by the function relating it to GDP per capita at MER (this ratio should not normally be very different from 1); (2) the current conversion ratio, which uses the potential GDP per capita (actual GDP per capita having not yet been computed) to feed the analytical function.

$$GDPP_r = GDP_r * PPPConV_r$$

$$GDPPCP_r = GDPPC_r * PPPConV_r$$

where

$$PPPConV_{r,t=1} = \frac{GDPPCP_{r,t=1}}{GDPPC_{r,t=1}}$$

$$PPPConV_{r,t} = PPPConV_{r,t-1} - (PPPConV_{r,t-1} - (FunctionRatioInitial_{r,t=1} * ConversionRatioCurrent_{r,t}))$$

$$FunctionRatioInitial_{r,t=1} = \frac{GDPPCP_{r,t=1}}{AnalFunc(GDPPC_{r,t=1})}$$

$$ConversionRatioCurrent_{r,t} = \frac{AnalFunc(GDPPOTPC_{r,t})}{GDPPOTPC_{r,t}}$$

Although the IFs model represents prices in real terms (no monetary sector and inflation), there are relative sectoral price changes and a separate section of the documentation describes the computation of a relative-price adjusted value added (RPA). The sum of those can also produce a relative-price adjusted GDP (GDPRPA). The utility of this variable can be seen, for instance, in countries, like Saudi Arabia, that are heavily dependent on energy exports. Large swings in the relative price of the energy sector can dramatically affect the relative-price adjusted GDP and we therefore carry this variable forward from the goods and services market into the more encompassing social accounting matrix representation.

$$GDPRPA_r = \sum^S VADDRPA_{r,s}$$

3.2.7 Trade

Imports and exports respond to regional prices (PRI) relative to those elsewhere. As a measure of prices elsewhere, IFs uses a world average price (WP), weighted by regional gross production and adjusted by the exchange rate (EXRATE).

$$WP_s = \frac{\sum^R (PRI_{r,s,t-1} * EXRATE_r * ZS_{r,s})}{\sum^R ZS_{r,s}}$$

The computation uses lagged prices because at this point the recursive equation system has not yet computed current prices.

Exports are responsive to both changes in production and changes in prices via respective elasticities. On the production side we begin by computing an export base (XBASE). It is initial exports (XS) plus some portion (XKAV) of growth in production, represented by potential gross sectoral production (ZSP). Representing a moving average of incremental production that is exported helps maintain global balance and stable behavior in long-term forecasting. The exported portion is modified over time in response to two elasticities. The first is a fairly standard elasticity of exports with income/production growth (*elasxinc*). The second is an export shift parameter that one would normally use to represent scenarios about export promotion or constraint.

$$XBASE_{r,s} = (XS_{r,s,t-1} + (ZSP_{r,s} - ZSP_{r,s,t-1}) * XKAV_{r,s} * \mathbf{elasxinc}_r) * (1 + \mathbf{xshift}_r)$$

Given the export base, export capacity (XC) responds to the difference between local and global prices using an elasticity of trade with prices (*elastrpr*). In addition we use two parameters (first order and second order) to pursue dynamically a global trade balance over time (*elprx1* and *elprx2*). You may wish more detail on the adjustment mechanism.

$$XC_{r,s} = XBASE_{r,s} * \left(1 + \left(\frac{PRI_{r,s,t-1} * EXRATE_r - WP_s}{WP_s} \right)^{\mathbf{elxpr1}} * \left(1 + \left(\frac{PRI_{r,s,t-1} * EXRATE_r - WP_s}{WP_s} - \frac{PRI_{r,s,t-2} * EXRATE_r - WP_{r,t-1}}{WP_{r,t-1}} * \mathbf{elastrpr}_s \right)^{\mathbf{elxpr2}} \right)$$

Import capacities or demands (MD) are almost exactly analogous. Whereas exports are tied to production, however, imports are tied to a demand base (DBASE) made up of final demands plus gross sectoral production potential (ZSP).

$$DBASE_{r,s} = CS_{r,s} + GS_{r,s} + INVS_{r,s} + ZSP_{r,s}$$

Imports are responsive to both changes in income and changes in prices and respective elasticities. We use the demand base as the basis for the income term. Specifically, we create a basic import level (MBASE) that grows with the demand base, responsive to an income elasticity parameter for imports (*elasminc*).

$$MBASE_{r,s} = MS_{r,s,t-1} + (DBASE_{r,s} - DBASE_{r,s,t-1}) * MKAV_{r,s} * \mathbf{elasminc}_r$$

A moving average for imports (MKAV) as a portion of changes in the demand base helps maintain global balance and stable behavior over the long-term.

Given this income responsive import base, regional prices relative to global ones (again with first and second order terms), an elasticity of imports with prices (*elastrpr*), and parameters for dynamic change of imports (*elprm1* and *elprm2*) largely determine import demands. One last factor, however, is a domestic protection multiplier (*protecm*), which can cause the import price (MPRIC) to rise or fall relative to the world price.

$$MD_{r,s} = MBASE_{r,s} * \left(1 + \left(\frac{PRI_{r,s,t-1} * EXRATE_r - MPRIC_{r,s}}{MPRIC_{r,s}} * \mathbf{elastrpr}_s \right) \right)^{elmpr1} * \left(1 + \left(\frac{PRI_{r,s,t-1} * EXRATE_r - MPRIC_{r,s}}{MPRIC_{r,s}} - \frac{PRI_{r,s,t-2} * EXRATE_{r,t-1} - MPRIC_{r,s,t-1}}{MPRIC_{r,s,t-1}} * \mathbf{elastrpr}_s \right) \right)^{elmpr2}$$

where

$$MPRIC_{r,s} = WP_s * \mathbf{protecm}_r$$

World export capacity (WXC) and world import demand (WMD) are simply sums across regions.

$$WXC_s = \sum^R XC_{r,s}$$

$$WMD_s = \sum^R MD_{r,s}$$

These will always be somewhat different. Actual world trade (WT) is the average of the two.

$$WT_s = \frac{WXC_s + WMD_s}{2}$$

We are now able to compute actual sectoral exports (XS) and imports (MS), normalizing the capacity for exports and the demand for imports to the actual world trade.

$$XS_{r,s} = WT_s \frac{XC_{r,s}}{WXC_s}$$

$$MS_{r,s} = WT_s \frac{MD_{r,s}}{WMD_s}$$

We can now update the moving average export (XKAV) and import (MKAV) propensities for the next cycle. This requires historical weights (*xhw* and *mhw*) for exports (XHW) and imports (MHW).

$$XKAV_{r,s,t+1} = XKAV_{r,s} * \mathbf{xhw} + (1 - \mathbf{xhw}) * \frac{XS_{r,s}}{ZSP_{r,s}}$$

$$MKAV_{r,s,t+1} = MKAV_{r,s} * mhw + (1 - mhw) * \frac{MS_{r,s}}{DBASE_{r,s}}$$

The above equations are necessary only for three of the five economic sectors. The agriculture and energy models compute trade in those sectors separately. Given the computation of sectoral exports and imports for all sectors (in this or other models), it is possible to compute total exports (X) and imports (M).

$$X_r = \sum^s XS_{r,s}$$

$$M_r = \sum^s MS_{r,s}$$

IFs also computes these in a "relative-priced adjusted" form, multiplying the real sectoral values by global prices. Why is this important? If, for example, a country were highly dependent on energy exports and the price of energy (from the energy model) doubled relative to other prices, failure to adjust trade for prices would understand the country's trade balance. It is the relative price-adjusted trade that is taken to the international financial calculations.

$$XRPA_r = \sum^s XS_{r,s} * WP_s$$

$$MRPA_r = \sum^s MS_{r,s} * WP_s$$

A purely optional adjustment to the relative price adjusted import and export levels is available for the model user who wishes to examine the hypothetical impact of changes in the terms of trade. The terms of trade parameter (*termtrm*) is a multiplier with a normal value of 1.0. Higher values shift the terms of trade in the favor of Southern or less-developed regions (those with initial GDP per capita less than an amount, such as \$5,000, specified by *nsdiv*) and lower values favor Northern regions. The determination of North and South by initial GDP per capita and the fixing of those categories is important.

if $GDPPC_{r,t=1} > nsdiv$ *then* $XRPA_r = XXRPA_r / termtrm$

$$MRPA_r = MRPA_r * termtrm$$

if $GDPPC_{r,t=1} \leq nsdiv$ *then* $XXRPA_r = XXRPA_r * termtrm$

$$MRPA_r = MRPA_r / termtrm$$

The adjustment to regional exports and imports for hypothetical terms-of-trade change can result in a failure of exports and imports to sum across regions to the same global total. At this point IFs thus normalizes regional exports and imports to a total value for

world trade equal to the average of the current sums of exports and imports across regions.

Given final values for regional exports and imports it is possible to compute regional trade balances (TRADEBAL), using relative-price adjusted exports and imports.

$$TRADEBAL_r = XRPA_r - MRPA_r$$

3.2.8 Stocks and Prices

IFs is fundamentally a general equilibrium model (GEM), but one in which inventory stocks serve as a temporary buffer between demand and supply and prices act to move the system towards equilibrium over time. The production available for final demand (PFD) and imports (MS) serve to increase stocks. Consumption (CS), investment (INVS), and government spending (GS) by sector of origin serve to decrease stocks, as do exports (XS).

$$ST_{r,s} = ST_{r,s,t-1} + PFD_{r,s} - CS_{r,s} - INVS_{r,s} - GS_{r,s} - XS_{r,s} + MS_{r,s}$$

Prices (PRI) in IFs are quite important in the agricultural and energy models where they directly affect demand and supply through elasticities. In the economic model they have lesser impact, primarily through trade and secondarily through price-responsiveness of sectoral consumption. In addition, prices implicitly affect investment by destination, although for reasons of computational sequence IFs actually uses stock levels directly to redirect investment by sector. Prices in IFs are relative prices and are indices around initial base values of "1." They are based on stock levels and a second order stock change term.

$$PRI_{r,s} = PRI_{r,s,t-1} * \left(1 + \frac{ST_{r,s} - dstl * STBASE_{r,s}}{STBASE_{r,s}} \right)^{elprst1} * \left(1 + \frac{ST_{r,s} - ST_{r,s,t-1} * GDP_{r,t=1}}{STBASE_{r,s}} \right)^{elprst2}$$

where

$$STBASE_{r,s} = ZSP_{r,s} + MS_{r,s,t=1} * \frac{GDP_{POT_r}}{GDP_{r,t=1}}$$

The stock base (STBASE) is the sum of gross production (important to large producers in an economic sector, whether they consume domestically or export) and initial imports scaled up by potential GDP growth (important to large importers, when domestic production is small). The desired level of stocks as a portion of the base (*dstl*) is exogenous.

3.2.9 Investment

The determination of investment by destination is a two-step procedure. First, IFs computes demand for investment by each sector (IFSDEM), responsive primarily to inventory levels. The base value of investment demand is dependent on the total level of gross capital formation (IGCF)—see its computation in the discussion of the social

accounting matrix—and the portion of that directed into a particular sector (IDK) during the last time step. Parameters (*elinst1* and *elinst2*) control the speed of adjustment. Second, IFs normalizes those demands to the total level of gross capital formation.

$$IFSDEM_{r,s} = IDK_{r,s,t-1} * IGCF_r * \left(1 + \frac{ST_{r,s} - dstl * STBASE_{r,s}}{STBASE_{r,s}} \right)^{elinst1} * \left(1 + \frac{ST_{r,s} - ST_{r,s,t-1}}{STBASE_{r,s}} \right)^{elinst2}$$

where

$$STBASE_{r,s} = ZSP_{r,s} + MS_{r,s,t=1} * \frac{GDPPOT_r}{GDP_{r,t=1}}$$

The above equation handles the sectors other than agriculture and energy. The model for agriculture provides investment need for that sector (IANEED), as does the energy model (IENEED).

$$IDSDEM_{r,s=1} = IANEED_r$$

$$IDSDEM_{r,s=2} = IENEED_r$$

To obtain actual investment by destination (IDS) we can distribute total investment (I) across sectors proportionately to their demands.

$$IDS_{r,s} = IGCF_r * \frac{IDSDEM_{r,s}}{\sum^s IDSDEM_{r,s}}$$

As an indicator (and as the basis for sectoral investment splits in the next time cycle) we can compute the fractions going to each sector (IDK).

$$IDK_{r,s} = \frac{IDS_{r,s}}{IGCF_r}$$

Capital stock (KS) in the next time period is simply the old capital stock plus investment by destination, minus depreciation (the capital stock divided by its lifetime, *lks*), and minus the portion of capital destroyed as civilian damage (CIVDM) in war (see the policy model). The last term is, of course, normally zero.

$$KS_{r,s,t+1} = KS_{r,s} + IDS_{r,s} + \frac{KS_{r,s}}{lks} - KS_{r,s} * CIVDM_r$$

IFs makes no effort to represent a gestation period for capital of more than 1 year. Although it would be desirable to do so, it would also require a "look ahead" capability of the model to plan capital requirements several years in the future. Such a feature would add some realistic cyclical behavior to the model, but would also be somewhat difficult to control. And it is not the aim of IFs to capture business cycles.

3.2.10 Indicators

The economic model computes several indicators of interest to many IFs users. These include global product (WGDP), global product per capita (WGDPPC), an absolute measure of the per capita GDP gap between more developed (D) and less economically developed (L) regions (NSGAPA), and a relative measure of the same gap (NSGAPR). A parameter (*nsdiv*) establishes the dividing point in thousand constant dollars between more and less developed.

$$WGDP = \sum^R GDP_r$$

$$WGDPPC = \frac{WGDP}{WPOP}$$

$$NSGAPA = \left(\frac{\sum^D GDP_r}{\sum^D POP_r} \right) - \left(\frac{\sum^L GDP_r}{\sum^L POP_r} \right)$$

$$NSGAPR = \frac{\left(\frac{\sum^D GDP_r}{\sum^D POP_r} \right)}{\left(\frac{\sum^L GDP_r}{\sum^L POP_r} \right)}$$

In an algorithmic procedure the system also computes a ratio of the GDP per capita (MER) in the countries that by average GDP per capita make up the richest 10 percent of global population to those that populate the bottom 10 percent (NSGAPRTB).

3.3 The Social Accounting Matrix (SAM)

3.3.1 Data for Initial Conditions in the Pre-processor

Finding data for all counties in IFs to initialize the financial flows of the social accounting matrices with the pre-processor (Hughes and Hossain 2003) is not just challenging, but impossible. IFs draws on data from a variety of sources, including the World Bank's World Development Indicators, the OECD, and the Global Trade and Analysis Project. The IMF's Government Finance Statistics (GFS are at <http://www.imf.org/external/data.htm>) provides a number of important series including the ability to begin differentiating central and local government revenues and expenditures, important in IFs as we increasingly represent states, provinces or other jurisdictions within countries. At this writing, most analysis is done at the general government level (central plus local) and local variables are rather preliminary and the distinction is primarily made using data from the IMF's GFS.

The pre-processor reads total government revenue from the SeriesGovtCalcRevTot%GDP table of the IFsHistSeries.mdb data file and reads central government revenue from SeriesGovtCalcRevCen%GDP. It computes the local portion as the residual. Whenever values are missing in tables for these variables or others in the

pre-processor, holes are filled for countries with values generally estimated from a cross-sectional relationship with GDP per capita at PPP.

Various revenue streams for general government are initialized in the preprocessor with data from the World Development Indicators or the IMF's GFS. These include taxes from firms (SeriesTaxCorp%Tot) from the IMF, social security and welfare taxes (TaxSocSec%CurRev) from WDI, and indirect taxes (TaxGoodServ%CurRev) from the WDI and originally the IMF GFS. Because many countries do not have direct taxes on households and data are poor, they are set as residual for revenue generation

Similarly, the pre-processor reads total government expenditures from SeriesGovtCalcExpendTot%GDP and central expenditures from SeriesGovtCalcExpendCen%GDP. When both expenditure values exist in the data, local is again the residual. If total expenditures do not exist, then local is computed based on some data about the growth of local revenues and expenditures in North, Walli and Weingast (2009: 10):

$$GOVEXPLOCAL_{r,t=1} = 2.0 + 12.0 * Amin\left(\frac{GDPPCP_{r,t=1}}{20}, 20\right)$$

In this case, total expenditures are initialized as the sum of central and local.

Expenditures break down into direct government consumption (GOVCON) in categories of military, health, education, infrastructure, R&D, and other and into government to household transfers (GOVHHTRN). Data for the various categories come primarily from WDI sources and GOVCON is set as the larger of the sum from the categories and a direct specification of it in the WDI data (SeriesGovCon%GDP). The health and education expenditures are assigned to GOVCONLOCAL, but GOVCON itself remains general government (central plus local). The pre-processor assures that GOVEXP is at least 1 percent more of GDP than is GOVCON, to make some headroom for government transfers.

Data on government transfers are again not very strong. We identify social security/welfare payments (GOVSSWEL) and pension payments (GOVHHPenT) as the two categories, based on WDI data in SeriesGovSSWelBen%Exp and SeriesGovtPension%GDP, respectively.

3.3.1.1 A note on forecasting central and local government finance

The process described above initializes four local government finance variables for a country as a whole, namely

GOVREVLOCAL (local government revenues)

GOVEXPLOCAL (local government expenditures)

GOVCONLOCAL (local government direct consumption)

GDSLOCAL for categories health and education (local government consumption by type)

In addition three variables have been added as placeholders that would add complexity and possible utility to representation of government finance at the sub-national level:

GOVREVCENFRLOCAL (government revenue at the central level from the local—transfers)

GOVREVLOCALFRCEN (government revenue at the local level from the central—transfers)

GOVREVLOCALTAX (government revenue raised through local taxes)

3.3.1.2 Possible futures steps in forecasting central/local government finance

At this point no dynamics have been added to the model for the four initialized variables, so they are displayed as constant across time. It would be easy to add dynamics that made them constant shares of GOVREV, GOVEXP, GOVCON, and GDS, respectively. It would take very considerably more work to make truly dynamic the central and local shares. There are two very different circumstances for such representations:

1. In integrated local representation without regionalization of countries. Steps of complexity: (a) local shares could be raised or lowered with functional forms or exogenous multipliers, with no forward linkages in model version with no specific representation of sub-units, having no real model impact; (b) passing through local changes to total revenues or expenditures as straight increases or decreases; for instance, if total expenditures were pushed up by an exogenous increase in local expenditures, it could have broad impacts on revenues as well as on the targets of those expenditures. But it is not clear that (b) would add much model value because exogenous changes to total revenues and expenditures are already possible.

2. In regionalized (states, provinces, districts) local representation. Using South Africa as an example, this is heavily complicated by the fact that the model cannot now represent both central government (e.g. national level government in Pretoria) and local government finance (e.g. in Western Cape), much less the interaction of the two and the possible implications for other provinces. The total government numbers being brought into the model for all SA provinces represent values related to provincial-level initialization.

A first step might be (a) to provide initialization data for as many as possible of the 7 variables for SA provinces and make sure the preprocessor uses these when available instead of estimates; (b) to initialize total government for all SA provinces as mirrors of SA national government, making all provincial total variables essentially identical shares of the national total percentage rates. The two steps are complicated, but reasonably limited.

A second and considerably larger step might be (c) to allow manipulation of provincial local expenditures with pass through to total (actually the provincial shares of the total) as indicated in point 1 above.

A much bigger step would be (d) to think about trade-offs and interactions across expenditures and revenues within and across local and total levels. It would then be possible, for example, to push up a specific consumption expenditure (say that on education) and boost educational performance and downstream impacts of that within one region, representing the trade-offs with (i) other expenditures in that region (this can already be done, but not always with meaningful provincial financial numbers) (ii) offsetting revenue increases at the local level (assuming there is any local sourcing of revenue); (iii) transfers from the regional total representation to or from the local; (iv) some combination since there would be many possible ones. In general, representing trade-offs outside of the specific region would be especially messy.

Given the lack of meaningful forecasting of local government finance variables at this time, the discussion on the forecasting of government finance will discuss only total government variables.

3.3.2 *Government Revenue*

The model represents five revenue streams from taxes on incomes on household incomes before taxes and transfers (HHIncBTT) and firm income before taxes and transfers (FirmIncBTT), which are the labor and capital shares of value added, respectively. Those streams are: household income taxes (HHTAX), household social security/welfare taxes (HHGOVSS), firm income taxes (FIRMTAX), firm social security/welfare taxes (FIRMGOVSS), and indirect taxes (INDIRECTTAX). We have no database to distinguish social security/welfare taxes assessed on households and firms. In each revenue equation there are exogenous multipliers that the user can manipulate to simulate changes in revenue streams (*hhtaxrm*, etc.) and total government revenues (*govrevm*). In addition there is a multiplier on revenues (*MulRev*) whose calculation we can see in the discussion of balancing of revenues and expenditures, and whose application lags one year behind the calculation. Some governments also have net domestic revenues from state-owned enterprises, but the model does not represent this stream (which is, in any case, often actually negative).

$$HHTAX_{r,h} = HHIncBTT_{r,h} * \frac{HHTAXR_{r,h}}{100} * hhtaxrm_r * govrevm_r * MulRev_{r,t-1}$$

$$HHGOVSS_{r,h} = HHIncBTT_{r,h} * \frac{SSWELTAXR_r}{100} * ssweltaxrm_r * govrevm_r * MulRev_{r,t-1}$$

$$FIRMTAX_r = FirmIncBTT_r * \frac{FIRMTAXR_r}{100} * firmtaxrm_r * MulRev_{r,t-1}$$

$$FIRMGOVSS_r = FirmIncBTT_r * \frac{SSWELTAXR_r}{100} * ssweltaxrm_r * govrevm_r * MulRev_{r,t-1}$$

$$INDIRECTTAX_r = FirmIncBTT_r * \frac{INDIRECTTAXR_r}{100} * indirecttaxrm_r * MulRev_{r,t-1}$$

Total domestic government revenues, before consideration of external revenues in the form of foreign aid (GovRevBA) is computed from the five streams, after summing the social security or welfare taxes across households and firms (SSWELTAX).

$$SSWELTAX_r = \sum^H HHGOVSS_{r,h} + FIRMGOVSS_r$$

$$GovRevBA_r = \sum^H HHTAX_{r,h} + FIRMTAX_r + SSWELTAX_r + INDIRECTTAX_r$$

It is useful also to compute the overall tax rate as an output indicator.

$$TAXRA_r = GovRevBA_r / GDP_r$$

For aid recipients only, the amount of government receipts adjusts government revenues.

if $AID_r > 0$ then

$$GOVREV_r = GovRevBA_r + AID_r$$

There are no dynamics yet in place for local government expenditures (GOVREVLOCAL). The model now has placeholder variables for net revenue transfers to local governments from central ones (GOVREVLOCALFRcen) and the direct imposition of taxes at the local level (GOVREVLOCALTAX), but they are now set at 0.0 because they are not yet supported by data in the model and no computations drive or use them.

3.3.3 Government Expenditure

In years beyond the base year the total of government expenditures is calculated from the sum of direct consumption and transfers. The two components, however, each require a moderately complex calculation. There are no dynamics yet in place within IFs for local government expenditures (GOVEXPLOCAL).

$$GOVEXP_r = GOVCON_r + GOVHHTRN_r$$

Government Consumption. Computation of government consumption (direct expenditures on the military, education, health, R&D, foreign aid, and other categories) begins with use of a function to compute estimated government consumption (EstGovtConsum) as a portion of GDP, using GDP per capita (PPP) as the driver. The function supports a behavioral assumption of generally increasing expenditures with increases in GDP per capita (although the rate of increase is very slow).

$$EstGovtConsum_r = AnalFunc(GDPPCP_r)$$

The estimated value then enters a convergence calculation that IFs uses in a number of instances. In the first year a ratio term (GovConR) was computed that represented the degree to which a country's consumption/GDP differed from the estimated value. That ratio multiplies the estimated term in future years, allowing the function normally to increase consumption/GDP as GDP per capita rises. At the same time, such divergence from estimated functions is almost as often a matter of data inadequacy or of temporary factors for a country as it is of persistent idiosyncrasy. The convergence function allows the country/region's value to converge towards the functional calculation over a period of time (*govfinconv*), usually quite long. Such convergence also helps avoid ceiling effects (e.g. government consumption as 100% of GDP) as GDP per capita rises.

The second term in the equation below is called the Wagner term, after the discoverer of the long-term behavioral tendency for government consumption to rise as a share of GDP, even at stable levels of GDP per capita. This is built into the consumption calculation through an exogenous parameter (*wagnerc*) that is multiplied by the number of the forecast year (at the time of this writing the coefficient's value was 0). The result of the convergence process and the multiplicative terms is divided by 100 and multiplied by a GDPBase term. The GDP base term could just be GDP. But relative price adjustments in countries can greatly affect government consumption—for instance, rising and falling relative oil prices can sharply affect government revenues and expenditures in oil-rich countries like those of the Gulf. Hence we arbitrarily add half of the difference between the GDP relative price adjusted value (GDPRPA) and the GDP to the base from which we compute GOVCON.

$$GOVCON_r = (Converge(EstGovtGonsum_r * GovConR_{r,t=1}, EstGovtCosum_r, govfinconv) * WagnerTerm * govexpm_r * MulExp_{r,t-1}) / 100 * GDPBase_r + WBFlowTerm_r + AidFlowTerm_r$$

where

$$GovConR_{r,t=1} = \frac{GOVCON_{r,t=1}}{\left(\frac{EstGovtCosum_{r,t=1}}{100} * GDP_{r,t=1}\right)}$$

$$WagnerTerm = 1 + t * wagnerc$$

$$GDPBase_r = GDP_r + (GDPRPA_r - GDP_r) * 0.5$$

$$WBFlowTerm_r = (XWBFLows_{r,ft=1,t} + XWBFLows_{r,ft=2,t}) - (XWBFLows_{r,ft=1,t=1} + XWBFLows_{r,ft=2,t=1})$$

$$AidFlowTerm_r = (AID_r / GDP_r) - (AID_{r,t=1} / GDP_{r,t=1}) * GDP_r * GovConPortion_r$$

Almost finally, government consumption is further modified by an exogenous multiplier of government expenditures (*govexpm*), allowing the user to directly control it by country/region and by an endogenously computed multiplier on expenditures (MulExp) that, parallel to MulRev, reflects the balance or imbalance in government expenditures

and the debt level (see the discussion of equilibration of revenues and expenditures for detail on its computation).

Finally, there are adjustments to reflect the possible impact of financial flows from international financial institutions (IFIs) such as the World Bank or International Monetary Fund (called XWBFLOWS independent of the source) and from foreign aid (AID). The IFI term represents the absolute difference between those flows in the current and initial years. The targets of those IFI flows are set exogenously via a parameter (*xwbsector*) that is dimensioned by target of flow (ft), namely education, health, unskilled households, skilled households, and other. Funds for the first two targets go to government consumption and funds for the household targets go to transfer payments. Extensive documentation of the stocks and flows associated with IFI funds and the user control of them is provided elsewhere. The aid flow term takes into account the difference between current and initial net receipts, allocating a portion of this to government consumption (the rest will go to household transfers). The government consumption portion is proportional to the share of government consumption in the sum of government consumption and household transfers.

As an alternative to the structure described above, the project has experimented with the computation of GOVCON not from a single function, but as the sum of estimates of its various destination components such as military and education (the GDS vector discussed below), with each of those estimates computed as a function of GDP per capita). That is rather more complex than necessary; and the model's currently quite rapid drop of infrastructure spending as a portion of GDP in the transition of low-income countries to middle-income ones tends to bring down or hold down the ratio of GOVCON to GDP more than historical data suggest occurs with development.

The division of government expenditures into target destination categories (GDS), part of the broader socio-political module of IFs is described later. That division is, of course, also a key agent-class behavior. With respect to sector of origin for government consumption (GS), which is information needed for the equilibration mechanism in the core commodities module, IFs simplistically assumes in the pre-processor (because of an absence of data or even qualitative information) that all government spending except arms, which has its source in manufactures, comes from services. On a year-to-year basis, the sectors of origin remained fixed at the initial proportions.

$$GS_{r,s,t} = GOVCON_{r,t} * \frac{GS_{r,s,t=1}}{GOVCON_{r,t=1}}$$

Government to Household Transfers. Government transfers, as distinguished from direct consumption expenditures, are computed using two different behavioral logics, a top-down one like the one for government consumption, and a bottom-up logic. The bottom-up logic is especially important in the analysis of pensions, because it is responsive to the changing size of the elderly population.

The top-down logic again uses an aggregate function responsive to changes in GDP per capita to calculate an estimated value of transfers as a portion of total government expenditures (EstGovtHHTrn).

$$EstGovtHHTrn_r = AnalFunc(GDPPCP_r)$$

The estimated value for the portion of total expenditures that household transfers constitute, adjusted by an initial shift term to accommodate initial data (GovHHTrnR), converges over time to the estimated value without the shift term. As with government consumption there is a Wagner term to represent possible growth with time, an exogenous multiplier for scenario analysis (*govexpm*) and an endogenously computed multiplier term to adjust government revenues and expenditures to each other (MulExp). That adjusted value is multiplied by an expected value for total government expenditures based on that in the preceding year, because the current value is not yet available.

$$GovHHTrnTop_r = (Converge(EstGovtHHTrn_r * GovHHTrnR_{r,t=1}, EstGovtHHTrn_r, govfinconv) * WagnerTerm * govexpm_r * MulExp_{r,t-1}) / 100 * EGovExp_r + AidFlowTerm_r$$

where

$$WagnerTerm = 1 + t * wagnerc$$

$$GovHHTmR_{r,t=1} = \frac{GovtHHTRN_{r,t=1}}{\left(\frac{EstGovtHHTRN_{r,t=1} * GOVEXP_{r,t=1}}{100}\right)}$$

$$EGovExp_r = GOVEXP_{r,t-1} * (1 \mp GDPGR_{r,t=1})$$

$$AidFlowTerm_r = (AID_r / GDP_r) - (AID_{r,t=1} / GDP_{r,t=1}) * GDP_r * GovTranPortion_r$$

Finally, there are adjustments to reflect the possible impact of financial flows from foreign aid (AID). The aid flow term takes into account the difference between current and initial net receipts, allocating a portion of this to government consumption (the rest goes to household transfers). The government transfer portion is proportional to the share of household transfers in the sum of government consumption and household transfers.

The bottom-up logic involves computing two terms, one for pensions and one for social welfare, in preparation for summing them. The pension term, like the top-down logic, also computes an initial estimate of pension expenditures (GovtPensions) as a portion of GDP using a function estimated cross-sectionally with GDP per capita at PPP. Not shown is an algorithm that ramps up that portion from 2 percent to 8 percent as GDP per capita at PPP climbs from 0 to \$8,000, over-riding the bottom portion of the analytical function below.

$$EstGovtHHTrnPen_r = AnalFunc(GDPPCP_{r,t})$$

An initial estimate of government spending on pensions (GovtPensions) is calculated by multiplying this estimate, adjusted for the shift between estimate and actual value in the first year, by an adjustment portion by a potential relative price adjusted GDP (GDPPOTRPA).

$$GovtPensions_r = EstGovHHTrnPen_r * GovHHTrnPenR_{r,t=1} * GDPPOTRPA_r$$

where

$$GovHHTrnPenR_{r,t=1} = \frac{GovtHHPEN_{r,t=1}}{\left(\frac{EstGovtHHPEN_{r,t=1} * GDP_{r,t=1}}{100}\right)}$$

Consistent with the bottom-up logic, that initial calculation is compared with a calculation that is built up from an estimate of the financial needs/demands of retired population. That second estimate looks at the size of the retired population, and multiplies that by per capita pension benefits in the first year, adjusted for the increase in GDP per capita. Given changing and often decreasing retirement ages, the retired population is a sum of those over age 65 and a term called RetirementPopAdj, which is computed to take into account an exogenous retirement age (*labretagem*). The larger of the analytically determined initial estimate and the retirement-population based numbers indicates the total pressure in the system for public pensions.

$$GovHHTrnPenBottom_r = AMAX(GovtPensions_r, (PopGT65_r + RetirementPopAdj_r) * GovHHPenPC_{r,t=1} * \frac{GDRPAPC_{r,t}}{GDRPAPC_{r,t=1}} MulExp_{r,t-1})$$

where

$$RetirementPopAdj_r = POPWORKING_r * (1 - labretagem_r)$$

A further algorithmic check is made of the above calculation to limit per-person pension requirements as the size of the retired population grows from below 20 percent of the total population (in which case the limit is the average per capita GDP) to 40 percent or more of the total (in which case the limit is 70 percent of the average per capita GDP).

The bottom-up calculation of welfare transfers is parallel to that for pensions.

$$GovHHTrnWelBottom_r = Pop_r * GovHHWelPC_{r,t=1} * \frac{GDPPC_{r,t}}{GDPPC_{r,t=1}} MulExp_{r,t-1}$$

The total bottom-up transfer calculation is the sum of that for pension and welfare transfers.

$$GovHHTrnTotalBottom_r = GovHHTrnPenBottom_r + GovHHTrnWelBottom_r$$

The larger of the two numbers indicates the total pressure in the system for transfers to households.

$$GOVHHTRN_r = AMAX(GovHHTrnTop_r, GovHHTrnTotalBottom_r)$$

Given these total transfers, there are further steps required to finalize the division into pension and welfare transfers and to divide each of those into transfers to skilled and unskilled households. The first step is an interim division of the total into pension and welfare categories using the bottom-up calculations.

$$GovHHPenTot_r = GOVHHTRN_r * \frac{GovHHTrnPenBottom_r}{GovHHTrnTotalBottom_r}$$

$$GovHHWelTot_r = GOVHHTRN_r - GovHHPenTot_r$$

The split to unskilled and skilled households is proportional to the share of their respective income in total household income and can be further affected by exogenous transfer parameters. The pension amounts are further adjusted by a pension multiplier (MulPen) that reflects the ratio of retirement needs (HHRETIRENEED) to life cycle consumption of those who are retired (CRETIRE) and helps equilibrate those two. The welfare amounts are further adjusted if net flows from international financial institutions (XWBFLOWS) have changed from initial conditions; those are added to the welfare transfers; welfare transfers are bounded to be below 20 percent of GDP.

$$GOVHHTTRNPEN_{r,h} = GovHHPenTot_r * \frac{HHINC_{r,h}}{\sum^H HHINC_{r,h}} * \mathbf{govhhtrnpenm}_{r,h} * MulPen_r$$

$$GOVHHTTRNWEL_{r,h} = GovHHWelTot_r * \frac{HHINC_{r,h}}{\sum^H HHINC_{r,h}} * \mathbf{govhhtrnwelm}_{r,h} + (XWBFLOWS_{r,ft=3 \text{ or } 4,t} - XWBFLOWS_{r,ft=3 \text{ or } 4,t=1})$$

where

$$MulPen_r = MulPen_{r,t-1} * \left(\frac{HHRETIRENEED_{r,t-1}}{CRETIRE_{r,t-1}} \right)^{0.1}$$

Finally, total government to household transfers are recalculated as the sum of the pension and welfare transfers to the unskilled and skilled households.

$$GOVHHTRN_r = GOVHHTTRNPEN_{r,h=1} + GOVHHTTRNPEN_{r,h=2} + GOVHHTTRNWEL_{r,h=1} + GOVHHTTRNWEL_{r,h=2}$$

3.3.4 Government Consumption by Destination

Government expenditures in IFs fall into two categories: direct consumption (GOVCON) and transfers (GOVHHTRN). The direct expenditures fall the following categories: military, health, education, research and development, infrastructure (the core types elaborated in the infrastructure model), other infrastructure (all non-specified infrastructure types), and other (including administrative and implicitly foreign aid of

donor countries). IFs divides total government consumption (GOVCON) into these destination sectors (GDS) in a process described below. The user can change that default pattern of government spending over time with a multiplier parameter (*gdsm*). The model normalizes the allocation to assure that the money spent is no more or less than total government consumption.

There are several issues that complicate the calculation of GDS, while simultaneously assuring that the sum of it across all categories equals GOVCON. These issues are:

1. In the cases of education, health, and infrastructure (core types), IFs contains models that produce calculations of expenditure needs or demands. Attempts to satisfy these demands in whole or part require adjustments in GDS for other spending categories.
2. In the cases of health, infrastructure and R&D categories, the model represents private as well as public spending (it should eventually also do this for education, but does not at this time). These private expenditures need tracking also and can, as in health, affect the ability of the government to meet total bottom-up health cost demands.
3. In the case of military spending, the model uses an action-reaction formulation that links spending of some countries to that of others (see the documentation of international politics). This is the reason that IFs code for GDS calculation was initially placed into and sits in the international political model (it should be moved to a separate function at some point).
4. In the case of education and health, financial flows into or out of countries from international financial institutions (IFIs) such as the World Bank need to adjust the values of GDS.

Preliminary GDS Calculations (Top Down Except for Infrastructure). The first stage in the computation of government spending by category (GDS) is the computation of a ratio of government spending in each category to GDP (G_k). In the sequencing of the code, in both first and subsequent years the first steps of that stage are the use of cross-sectionally estimated analytical functions that compute expected military, education, health and other infrastructure expenditures as a percent of GDP as a function of GDP per capita at PPP (G_kComp). A similar analytical function is used to compute total (public and private) R&D expenditures. In the first year there are actually data on such expenditures, so a country-specific "shift" vector is computed (G_kRI) that saves the ratios of the actual values to those computed in the functions. In the subsequent years, the values in that shift vector converge to 1 over 200 years, meaning that the pattern generated by the expected values from the functions very gradually take over from the pattern in the initial conditions. In the first year the other category is available as a residual (the preprocessor keeps it from going to zero or negative) and the ratio of that value to GDP (G_k) is maintained as the expected value over time. Note that there is no function or process to compute an expected value for infrastructure spending as a portion of GDP per capita; the data for infrastructure led model developers to conclude that it

was better to use only bottom-up information on infrastructure costs and expenditures. So in general:

$$Gk_{r,g,t} = (\text{Converge}(GkRI_{r,g,t=1}, 1, 200)) * GkComp_{r,g,t}$$

where

$$GkComp_{r,g,t} = \text{AnalFunc}(GDPPCP_{r,t})$$

$$GkRI_{r,g,t=1} = \frac{GDS_{r,g,t=1}/GDP_{r,t=1}}{GkComp_{r,g,t=1}}$$

After the initial calculation of the ratio of government spending to GDP in each category, a number of adjustments are made that vary by expenditure category in order to get an initial estimate of actual expenditures (GDS). First, for military spending the estimate use the Gk variable, but modifies it by both an exogenous multiplier (***gdsm***) and a multiplier for it (GkMul) computed in the action-reaction process around military spending (described in documentation of the international political model).

$$GDS_{r,g=mil} = GDP_r * Gk_{r,g=mil} * \mathbf{gdsm}_{r,g=mil} * GkMul_r$$

For health, R&D, and other, the equation is a variant in which funds coming from international financial institutions (GDSWB) are added to the basic calculation; in reality, such funds affect GDS only for education and health and at this point they are added only for health (nothing augments R&D and Other).

$$GDS_{r,g=health,R\&D,other} = GDP_r * Gk_{r,g=health,R\&D,other} * \mathbf{gdsm}_{r,g=health,R\&D,other} + GDSWB_{r,g=health,R\&D,other}$$

For education the equation is the same, except for the exogenous multiplier. That multiplier is used only later in the process, when the initial GDS estimate is being reconciled with the bottom-up cost calculation from the education model.

$$GDS_{r,g=educ} = GDP_r * Gk_{r,g=educ} + GDSWB_{r,g=educ}$$

The process is slightly more complicated for R&D because there are both private and public expenditures to account for. The code first computes a ratio of the expected share of public expenditures within GDP in the current year relative to the first year, a term that will be used as a multiplier in the next step (RandDMul).

$$RandDMul_{r,t} = \frac{\text{AnalFunc}(GDPPCP_{r,t})}{\text{AnalFunc}(GDPPCP_{r,t=1})}$$

That term is used to adjust the initial ratio of public R&D expenditures as a portion of GDP (RandDI) in order to compute the current time step's value, which is allowed to

converge, but at minimal speed (1000 years) to the estimated value for the current time step:

$$Gk_{r,g=R\&D,t} = \text{Converge}(RandDGDP_{r,t}, \text{AnalFunc}(GDPPCP_{r,t}), 1000)$$

where

$$RandDGDP_{r,t} = GDS_{r,g=R\&D,t=1} / GDP_{r,t=1} * 100 * RandDMul_{r,t}$$

Finally, this allows the computation of the initial estimate of GDS for R&D:

$$GDS_{r,g=R\&D} = GDP_r * Gk_{r,g=R\&D} / 100 * \mathbf{gdsm}_{r,g=R\&D}$$

At this point, then, initial estimates of GDS are available for all government spending categories except core infrastructure, other infrastructure, and education.

For the two infrastructure terms the process depends on value of a parameter that turns on government finance for them (*infrafinon*) and the default value is for it to be on (=1). In that case, for other infrastructure the process is essentially the same as for the military or health spending, except that there are no exogenous multipliers or other terms affecting the basic calculation. That basic calculation is:

$$GDS_{r,g=InfraOther} = GDP_r * Gk_{r,g=InfraOther}$$

where

$$Gk_{r,g=InfraOther,t} = (\text{Converge}(GkRI_{r,g=InfraOther,t=1}, 1, 200)) * GkComp_{r,g=InfraOther,t}$$

$$GkComp_{r,g=InfraOther,t} = \text{AnalFunc}(GDPPCP_{r,t})$$

$$GkRI_{r,g=InfraOther,t=1} = \frac{GDS_{r,g=InfraOther,t=1} / GDP_{r,t=1}}{GkComp_{r,g=InfraOther,t=1}}$$

For core infrastructure, when the *infrafinon* parameter has its default value, the value of GDS comes completely from a function (CalcInfraBudgetDemand) that adds up the public portions of the new construction and maintenance costs for various types of infrastructure for the current year as already calculated in the infrastructure model. For reasons said above, infrastructure GDS calculation, unlike other GDS categories, do not use any reference to either initial values of infrastructure spending or any expected value relative to GDP per capita at PPP. It is the only preliminary GDS value that is not calculated using a top-down logic.

$$GDS_{r,g=Infra} = \text{CalcInfraBudgetDemand}_r$$

where;

$$\begin{aligned} \text{CalcInfraBudgetDemand}_r &= \sum_{infct} \text{infrainvnewpubshrm}_{r,infct} * \\ &\text{INFRAINVESTNEW}_{r,infct} + \sum_{infct} \text{infrainvmaintpubshrm}_{r,infct} * \\ &\text{INFRAINVESTMAINT}_{r,infct} \end{aligned}$$

In the equations above, *infct* stands for the index number of various infrastructure types included in the IFs model, for example, road, electricity, water connections etc.; *INFRAINVESTNEW* and *INFRAINVESTMAINT* are total costs of new construction and maintenance respectively and *infrainvnewpubshrm* and *infrainvmaintpubshem* are shares of the costs borne by the government.

Bottom-Up Cost Calculation. The equation just above used a bottom-up cost calculation from the infrastructure model to compute an initial estimate of GDS for core infrastructure. In the case of both education and health, the models also provide a bottom-up cost calculation foundation and we turn to that next before proceeding to reconciliation of the Top-Down and Bottom-Up values.

We begin with education. The basic logic of that cost calculation is that in the first year of the forecast a ratio (*EdCostGDSEdRI*) is calculated of the public expenditures on education (GDS from data and the preprocessor) and a variable from the education model representing bottom-up costs (*EdTotCost*). A modeling decision was made to make this ratio converge to 1.1 over time.

$$\begin{aligned} \text{EdCostGDSEdRI}_{r,t} \\ = 1.1 + \text{Converge}(GDS_{r,g=educ,t=1}/EDTOTCOST_{r,t=1} - 1.1, 0, 70) \end{aligned}$$

The bottom-up educational expenditure demand (*DemandCalc*) used in the budget reconciliation process below is calculated as the bottom-up calculation in the preceding year times this ratio. In addition, a small adjustment is made to the bottom-up calculation from the previous year in order to make it grow somewhat over time; the initial growth rate is set equal to the initial growth rate of GDP (*igdpgr*) and allowed to converge to 0 over time.

$$\text{DemandCalc}_r = \text{EDTOTCOST}_{r,t-1} * \text{EdCostGDSEdRI}_{r,t} * (1 + \text{EcGrTerm}_r)$$

where

$$\text{EcGrTerm}_r = \text{Converge}(\text{igdpgr}_{r,t=1}, 0, 50)$$

[This would be logical place to discuss bottom up health cost wrap around from previous year - at this point I would simply wrap it with a moving average growth term, not using anything like the 1.1 stuff above which I don't understand. Health might logically precede education given their sequence in the GDS matrix]

Reconciling Top-Down Expenditure Availability with Bottom-Up Cost Calculation.

At this point the initial values of GDS have been calculated and the bottom-up cost

calculations for health, education, and infrastructure have been identified. But the preliminary GDS values will almost certainly not sum to GOVCON, which is what the model has determined is available. Nor will the bottom-up (EDTOTCOST and HLCOST) and top-down estimates for costs of and expenditures on health and education be equivalent. The two reconciliation issues are interactive and must be addressed to some degree simultaneously.

For education the first step in the reconciliation process (focusing on the bottom-up and top-down aspects) is a recalculation of GDS for education that determines a refined value taking into account both the bottom-up demand calculation (DemandCalc) and the earlier preliminary top-down version of GDS. The parameter for budget balancing (*edbudgon*), with a default value of 0.4 and constrained to be between 0 and 1, provides a weighted average of the two input terms.

$$GDS_{r,g=educ} = DemandCalc_r * edbudgon + GDS_{r,g=educ} * (1 - edbudgon)$$

[problem with gdsnormalize and RANDDEXP]

Because the model computes the value of total (public and private) research and development expenditures (RANDDEXP) and the normalization process will change the public portion of that, the value of RANDDEXP is temporarily changed to remove the public expenditures. After GDS for R&D is recalculated below, the revised version of GDS for R&D will be added back into the total.

$$RANDDEXP_r = RANDDEXP_r - GDS_{r,g=R\&D} / GDP_r * 100$$

The next step in the process is to normalize the sum of all GDS terms across categories to GOVCON. This process has the potential, however, of setting GDS values for health, education, and infrastructure that are very different from the bottom-up costs. Thus the normalization process includes a mechanism for avoiding that. It protects some or all of the bottom-up calculations during the normalization.

The mechanism for protecting some or all of the bottom-up calculations involves "set asides". The extent of set asides for education, core infrastructure and other infrastructure is determined by parameters (*budbal* and *infrabudsdrat*) that take on values between 0 and 1; values of 0 protect none of the bottom-up value and values of 1 protect it all. The default value of *infrabudsdrat* is 0.

$$GDSEdSetAside_r = GDS_{r,g=educ} * edbudgon$$

$$GDSInfraSetAside_r = GDS_{r,g=infra} * infrabudsdrat$$

$$GDSInfraOtherSetAside_r = GDS_{r,g=infraother} * infrabudsdrat$$

There is another step before the normalization process begins, namely adjusting GDS for education, core infrastructure, and other infrastructure by an exogenous multiplier (*gdsm*), the default value for which is 1.0.

$$GDS_{r,g=educ,infra,otherinfra} = GDS_{r,g=educ,infra,otherinfra} * gds_{r,g=educ,infra,otherinfra}$$

The sum of all GDS values (GTOT) can now be compared with GOVCON. If there are more than enough funds to provide all of the GDS requested (that is, GOVCON exceeds GTOT), the final calculation of GDS simply assigns proportional shares of the surplus to each GDS category.

$$GDS_{r,g} = GDS_{r,g} * \frac{GOVCON_r}{GTOT_r}$$

If there is a shortage of funds to meet all demand (that is, GTOT exceeds GOVCON), it would be possible to proportionately reduce each GDS accordingly. But the set aside values allow protection of some or all of the GDS values for education, infrastructure and other infrastructure. In the normalization process when there is a shortage of funds, each potentially or actually protected GDS is reduced by the set aside, as are both GTOT and GOVCON (using a GOVCON_{Red} or GOVCON reduced). Then the normalization of all GDS occurs and the protected or set aside values are added back in.

$$GDS_{r,g=educ} = GDS_{r,g=educ} - GDSEdSetAside_r$$

$$GDS_{r,g=infra} * = GDS_{r,g=infra} * - GDSInfraSetAside_r$$

$$GDS_{r,g=infraother} = GDS_{r,g=infraother} - GDSInfraOtherSetAside_r$$

$$GTOT_r = GTOT_r - GDSEdSetAside_r - GDSInfraSetAside_r - GDSInfraOtherSetAside_r$$

$$GOVCONRed_r = GOVCON_r - GDSEdSetAside_r - GDSInfraSetAside_r - GDSInfraOtherSetAside_r$$

$$GDS_{r,g} = GDS_{r,g} * \frac{GTOT_r}{GOVCONRed_r}$$

$$GDS_{r,g=educ} = GDS_{r,g=educ} + GDSEdSetAside_r$$

$$GDS_{r,g=infra} * = GDS_{r,g=infra} * + GDSInfraSetAside_r$$

$$GDS_{r,g=infraother} = GDS_{r,g=infraother} + GDSInfraOtherSetAside_r$$

At this point it is possible to add the public R&D expenditures back into the variable RANDEXP from which they had earlier been removed (leaving that variable temporarily to hold only private expenditures).

$$RANDEXP_r = RANDEXP_r + GDS_{r,g=R\&D}/GDP_r * 100$$

There are a number of forward linkages from GDS that are important elsewhere in the

model. The most important of these are linkages to multifactor productivity from human capital. Those are discussed elsewhere in this documentation. Here we note only two such forward linkages, both of which are set up in the same section of code as the computation of GDS.

The first is of a mortality multiplier (MORTMG) that is computed for the demographic model, using changes in health spending from the initial year and a parameter of the impact of that spending (*elashc*). This multiplier is now not typically used in the IFs system because the calculation of mortality that it feeds in the population model has been replaced in the default model mode by the entire health model.

$$MORTMG_r = 1 - \left(\frac{GDS_{r,g=health}}{GDP_r} - \frac{GDS_{r,g=health,t=1}}{GDP_{r,t=1}} \right) * elashc$$

The second forward linkage at this point is in the calculation of a saved value of health expenditures as a portion of GDP to be used in the calculation of water and sanitation in the infrastructure model. In the formulation below, if the exogenous multiplier *gdsm* for health spending is greater than or equal to 1 (the default), the value of *sHlthExpPcntGDP* is not allowed to decline over time.

$$sHlthExpPcntGDP_r = \frac{GDS_{r,g=health}}{GDP_r} * 100$$

3.3.5 Government Balances and Dynamics

Given calculations of government revenue and expenditure, it is possible to compute the government balance, adjusted by foreign aid donations when given (for donors, the sign of AID is negative). Repayments of loans to the IMF or World Bank also reduce the balance.

$$GOVBAL_r = GOVREV_r - GOVEXP_r + AID_r - XIMFCRFOUT_r - XWBLNFOUT_r$$

That allows the update of absolute government debt and a calculation of its magnitude relative to GDP. We use the relative price adjusted GDP because it reflects changes in energy and agricultural prices that can be of considerable significance for countries with heavy dependence on trade in those sectors.

$$GovDebtAbs_r = GovDebtAbs_{r,t-1} - GOVBAL_r$$

$$GOVDEBT_r = GovDebtAbs_r / GDPRPA_r$$

It is government debt as a percentage of GDP that IFs uses to build equilibrating dynamics for government revenues and expenditures.

In years beyond the first, the government debt as a percent of GDP will give rise to pressure for higher or lower levels of government revenues and expenditures. Those pressures will be conveyed via two multipliers applied to calculations of taxing and

spending in subsequent years. Those multipliers are set at "1" in the first year, indicating no change in pressures initially. The rest of this section explains how those multipliers are changed over time.

$$MulRev_{r,t=1} = 1$$

$$MulExp_{r,t=1} = 1$$

The first step in feeding back government debt level to pressures for increases or decreases in revenues and expenditures is to compute a target for government debt as a percent of GDP with which to compare the actual level. A function relating typical levels of debt to GDP per capita is used as the foundation of that target. Dynamically, the target begins as the initial ratio of debt to potential, relative price adjusted GDP, but we found that it enhanced long-term model behavior to set the target as a moving average of that ratio over time, only allowing the moving average to be changed over time when the debt to GDP ratio falls (thus reducing the target). Further the target converges to the minimum of the moving average and the value from the function.

$$GovDebtTarget\%_r = Converge(GovDebtMA_r, AMin(GovDebtMA_r, GovDebtComp_r, 100))$$

where

$$GovDebtComp_r = AnalFunc(GDPPCP_{r,t})$$

and where if $\frac{GovDebtAbs_r}{GDPPOTRPA_r} < GovDebtMA_{r,t-1}$ then

$$GovDebtMA_r = GovDebtMA_r * 0.6 + \frac{GovDebtAbs_r}{GDPPOTRPA_r} * 100 * 0.4$$

The tricky part is to get a dynamic system to chase the target over time, adjusting revenues and expenditures annually as it does so. IFs does that in this instance and in others by using an adjustment to feedback parameters based on two terms: an integral term (the absolute distance of the system from the target) and a differential term (the change in values of the target relative to the preceding year). Engineers refer to this process as a PID controller. The two terms are computed as Diff1 and Diff2. A potential relative price adjusted GDP term is used rather than the actual GDP of the year in order to protect systemic stability over time (use of GDP in the target term can set up oscillations in some feedback loops also involving GDP).

$$Diff1_r = GovDebtAbs_{r,t} - GovDebtTarget\%_r * GDPPotRPA_r$$

$$Diff2_r = Diff1_r - Diff1_{r,t-1}$$

Once the two terms are available, the PID adjuster routine (ADJUSTR), described elsewhere is called by IFs to convert the difference terms, modified by exogenous parameters, into a multiplier on the cumulatively-computed revenue multiplier term used in equations above. (This function historically used the parameters *elgrevdebt1* and *elgrevdebt2*, but those are currently hard-coded in the equation below as 0.15 and 0.30).

$$Adjstr(GDPPot_r, Diff1, Diff2, elgrevdebt1, elgrevdebt2, Mul)$$

$$MulRev_r = MulRev_{r,t-1} * Mul$$

In the above specification, the adjuster uses two elasticities for the difference terms. In completely parallel fashion, adjustment is made to the expenditure multiplier. Tuning of the model reinforced the need for the elasticities on the expenditure side normally to be lower than those on the revenue side (governments are more likely to adjust revenues than expenditures). *Elgexdebt1* and *elgexpdebt2* were set at 0.05 and -0.1, respectively. Yet when revenues and expenditures get substantially out of balance, there are times that some "extra kick" on the expenditure adjustment is needed and that is provided by the elasticity multiplier (ElMul) based on the ratio of revenues and expenditures.

$$Adjstr(GDPPot_r, Diff1, Diff2, elgexpdebt1 * ElMul_r, elgexpdebt2 * ElMul_r, Mul)$$

$$MulExp_r = MulExp_{r,t-1} * Mul$$

where

$$ElMul_r = AMIN \left(AMAX \left(\frac{GOVEXP_r}{GOVREV_r}, \frac{GOVREV_r}{GOVEXP_r} \right), 5 \right)$$

3.3.6 Household and Firm Income

It is possible to compute household income (before taxes and transfers) by sector, and as important, to divide household income into skilled and unskilled categories. The labor coefficients (LaborF) computed from the GTAP database allow that; the IFs IO threshold system creates average versions across countries of input-output matrices and also labor skill levels at increasing levels of GDP per capita at PPP, and then uses GDPPCP to fill country-holes in initialization and to change the skill level profile as GDPPCP changes over time. Those sector-specific coefficients are weighted by sector-specific value added to compute the total relative-price-adjusted value added (VADDRPA) going to unskilled and skilled households, respectively (VADDHHU and VADDHHS). Those shares are then summed to provide household income for each household type (HHINC). Firm income is the remaining GDP (equivalent to the capital share), augmented by any foreign direct investment (XFDIFIN) or portfolio investment (XPORTFIN).

The income share computation also introduces a skill shift term (SkillShiftMul). The basis for the term is a presumption that over time there is a greater demand for skilled labor and a commensurately lower demand for unskilled labor. The rate of annual change

has been initially set at 0.3 to 0.4%, corresponding approximately to the core level of change in multifactor productivity (analysis suggests that the rate might be a little low). There is much room for endogenization and refinement of this highly aggregate and rough specification. The importance of it lies in the re-allocation of household income away from unskilled households to skilled households. Because of the great growth in education around the world, and therefore the proportionately faster growth of skilled households, there is, in fact, credential inflation; in the absence of it, the supply of skilled labor would be increasing much faster than the demand for it and the price of it relative to unskilled labor would drop sharply. That is not happening in the world.

$$VADDHHU_{r,s} = VADDRPA_{r,s} * LaborF_{r,h=1,s} / SkillShiftMul_r$$

$$VADDHHS_{r,s} = VADDRPA_{r,s} * LaborF_{r,h=2,s} / SkillShiftMul_r$$

where

$$LaborF_{r,h,s}$$

= $F(GTAP$ labor skill data with IO matrices, IFs IO threshold system, $GDPPCP_r)$

$SkillShiftMul_r = SkillShiftMul_{r,t-1}^{1.004 \text{ to } 1.003}$ where exponent decreases with $GDPPCP$

$$HHINCShr_{r,h=1} = \sum^S VADDHHU_{r,s}$$

$$HHINCShr_{r,h=2} = \sum^S VADDHHS_{r,s}$$

$$HHIncBTT_{r,h=1} = \frac{HHINCShr_{r,h=1}}{HHINCShr_{r,h=1} + HHINCShr_{r,h=2}} * GDP_r * (1 - CDALF_r)$$

$$HHIncBTT_{r,h=2} = \frac{HHINCShr_{r,h=2}}{HHINCShr_{r,h=1} + HHINCShr_{r,h=2}} * GDP_r * (1 - CDALF_r)$$

$$VADDF_{r,s} = VADDRPA_{r,s} - VADDHHU_{r,s} - VADDHHS_{r,s}$$

$$FIRMINC_r = \sum^S VADDF_{r,s} + XFDIFIN_r + XPORTFIN_r$$

.

We need also to know the net flow of funds from firms to households, and how that is split between the two types of households. A simplifying intermediate step is to compute firm income after taxes (see the discussions of government revenues and expenditures for explanations of terms such as $FIRMTAX$, $FIRMGVSS$, $INDIRECTAX$, $HHTAX$, $GOVHHTRNWEL$).

$$FirmIncomeAfterTaxes_r = FIRMINC_r - FIRMTAX_r - FIRMGOVSS_r - INDIRECTTAX_r$$

IFs computes in the first year a firm investment ratio (FirmInvRI) as a general estimate of the portion of firm income after taxes that is used for gross capital formation. The residual after capital formation (investment) is the portion that is passed back to households as dividends and interest. In the absence of data on the distribution of dividends and interests between unskilled and skilled labor-based households, equations arbitrarily assign the overwhelming share of it, namely 90%, to skilled households.

$$HHDivInt_{r,h=1} = FirmIncomeAfterTaxes_r * (1 - FirmInvRI_{r,t=1}) * .1$$

$$HHDivInt_{r,h=2} = FirmIncomeAfterTaxes_r * (1 - FirmInvRI_{r,t=1}) * .9$$

where

$$FirmInvRI_{r,t=1} = Amin(1, \frac{I_{r,t=1}}{FirmIncomeAfterTaxes_{r,t=1}})$$

At this point it is possible to compute household income adjusted by all transfer payments from government and firm's dividends. Household disposable income (HHDispInc) is that gross income minus taxes. It is disposable income that IFs takes to the calculation of consumption and, as a residual, net savings.

$$HHINC_{r,h} = HHIncBTT_{r,h} + GOVHHTRNWEL_{r,h} + GOVHHTRNWEL_{r,h} + HHDivInt_{r,h}$$

$$HHDispInc_{r,h} = HHINC_{r,h} - HHTAX_{r,h} - HHGOVSS_{r,h}$$

It is also desirable to know the household income per capita (HHINCPC), which requires knowing the size of the household population (HHPOP) in each household type (that is, skilled and unskilled). The model assigns total population to households based on the labor supply (LABSUP) in each household type, thereby assuming equal household sizes.

$$HHPOP_{r,h} = POP_r * \frac{LABSUP_{r,h}}{LAB_r}$$

$$HHINCPC_{r,h} = \frac{HHINC_{r,h}}{HHPOP_{r,h}}$$

3.3.7 Household Consumption and Net Savings

One of the key behavioral relationships for households (and a very important one for the larger model) is the division of income between household consumption, C, and household savings (HHSAV). Many factors affect that division including (1) an inertial element of consumption in the face of changing household income, (2) the relationship between GDP per capita and the propensity of households to consume from income (it

tends to drop slowly), (3) the age structure of the population (people tend to consume a smaller proportion of income in the mid working years of life), (4) interest rates (high rates divert some income to savings), (5) life expectancy (longer life expectancy beyond retirement age should lead increase savings propensity). This section will address each of these in turn.

It is, however, not fully possible to conceptualize household consumption independent of other expenditure components (government consumption, investment, and net trade). There are equilibrating mechanisms that cross over the goods and services market and influence agent behavior with respect to consumption, investment and savings. Other topics cover those interacting representations.

3.3.7.1 The permanent income foundation of consumption

Households have ongoing consumption needs and patterns. Thus their consumption tends not swing as widely as does income. For that reason we begin by computing an internal model variable defined in the literature (e.g. Pistaferri 2001) as permanent household income (HHDIsIncPerm). The computation uses a weighted average to smooth changes of permanent income in the face of changes in actual annual disposable income (HHDIsInc).

$$HHDIsIncPerm_{r,h} = HHDIsIncPerm_{r,h,t-1} * (1 + IGDPRCOR_r) * .7 + HDDIsInc_{r,h} * .3$$

The sum permanent income across household types yields total household permanent income (TotHHIncPerm).

3.3.7.2 Basic consumption computation and the affect of GDP per capita

The core of the consumption equation involves calculating a preliminary estimate of household consumption for each of the two household types (ConsumHHPreI) by multiplying a consumption ratio (CRA) as a portion of permanent disposable household income times that income. The consumption ratio is carried along from year-to-year, providing an additional inertial but moving element of consumption (we will update it after finalizing the computation of consumption).

$$ConsumHHPreI_{r,h} = HDDIsIncPerm_{r,h} * CRA_{r,h,t-1}$$

Our own cross-sectional analysis confirms a tendency for the consumption share of income to decline somewhat with increases in GDP per capita. We use that function to calculate the average household consumption propensity (HHCPop). The second term in the computation of consumption (C) adjusts the preliminary computation by household with the change of this propensity over time. The third term adjusts this consumption term based on change over time in the ratio (PenShrInc) of government to household pension transfers (GOVHHPEM) to the total permanent income, assuming that pensions are mostly consumed.

$$C_r = \sum^H \text{ConsumHHPre}_{r,h} * \frac{\text{HHCProp}_{r,t}}{\text{HHCProp}_{r,t=1}} + (\text{PenShrInc}_{r,t} - \text{PenShrInc}_{r,t=1}) * \text{TotHHIncPerm}_r$$

where

$$\text{HHCProp}_r = \text{AnalFunc}(\text{GDPPCP}_r)$$

$$\text{PenShrInc}_r = \frac{\text{GOVHHPENT}_r}{\text{TotHHIncPerm}_r}$$

Many countries have initial consumption shares of GDP that are especially high or low for a variety of historical reasons (high receipt of external assistance or remittances can boost C and governmentally-induced savings can reduce it). So we also set up a very slow convergence of C toward the more typical cross-sectional pattern for it, using again the household propensity to consume to drive C toward that target value (CTarget) in one of our standard IFs functions, *ConvergeOverTime*. The target is a portion of the potential, relative price adjusted GDP (GDPPOTRPA), again adjusted by changes in pensions.

$$C_r = \text{ConvergeOverTime}(C_r, \text{CTarget}_r, 150)$$

where

$$\text{CDTarget}_r = \frac{\text{HHCProp}_r}{100} * \text{GDPPOTRPA}_r + (\text{PenShrInc}_{r,t} - \text{PenShrInc}_{r,t=1}) * \text{TotHHIncPerm}_r$$

3.3.7.3 The impact of age structure on consumption

Modigliani (1976) sketched the pattern of life-cycle savings and consumption and many others have followed (e.g. Lee and Miller 1992, Zhang and Zhang 2005, Zhang and Zhang 2009, Lee and Mason 2011). Consumption as a portion of income tends to be lowest during the peak working years and higher for the young and the old and/or retired.

Drawing especially on Lee and Mason (2011) we structured a relationship between marginal propensity to consume and age (the percentage portion of additional household income that would typically be consumed at different ages). We use that to differentiate consumption demands or needs of the population prior to working years (CPREWORK), during working years (CWORKING) and in retirement (CRETIRE). The function is used to compute an estimate of the percentage of additional income that each of the three groups consumes (YouthCon, WorkerCon, and PensionCon).

$$\text{YouthCon}_r = \text{AnalFunc}(\text{workageentry}_r / 2)$$

$$\text{WorkerCon}_r = \text{AnalFunc}((\text{workageentry}_r + \text{workagretire}_r) / 2)$$

$$\text{PensionCon}_r = \text{AnalFunc}(\text{workagretire}_r + \text{LIVEXP}_{r,t-1}) / 2)$$

At this point there is a potential diversion from the normal flow of computation is the parameter *cpctnduse* (C percent response to balance between the savings need and expected availability for use) has been assigned a value larger than its normal 0.0. If the value is positive, SavingsNew and SavingsOld terms are calculated using an annuity formula ($PV = C * [(1-(1+i)^{-n})/i]$) that is meant to assure adequate retirement income when C is set to 20.

$$SavingsNew_r = 20 * \left(\left(1 - (1 + savdingsdiscountr_t_r/100)^{\frac{workageentry_r - WORKAGERETIREND_r}{savingsdiscountr_t_r/100}} \right) \right)$$

$$SavingsOld_r = 20 * \left(\left(1 - (1 + savdingsdiscountr_t_r/100)^{\frac{workageentry_r - workageretire_r}{savingsdiscountr_t_r/100}} \right) \right)$$

The difference between the new and old savings suggests the additional (or lesser) savings needed.

$$SavingsNeeded_r = SavingsNew_r - SavingsOld_r$$

A portion of that (determined by the same optional parameter) can be subtracted from the consumption of workers so that they will instead save more.

$$workercon_r = workercon_r - SavingsNeeded_r * cpctnduse_r$$

Returning to the flow of computations regardless of the value of *cpctnduse*, we also compute the share of the three groups in the total population.

$$YouthPopShr_r = \frac{POPPREWORK_r}{POP_r}$$

$$WorkerPopShr_r = \frac{POPWORKING_r}{POP_r}$$

$$PensonerPopShr_r = \frac{POPRETIRED_r}{POP_r}$$

That allows an estimate of the share that each group would have of the total household consumption.

$$YouthConShr_r = (YouthPopShr_r * YouthCon_r)/100$$

$$WorkerConShr_r = (WorkerPopShr_r * WorkerCon_r)/100$$

$$PensionerConShr_r = (PensionerConShr_r * PensionerConShr_r)/100$$

The sum of the consumption shares follows.

$$ComputedCon_r = YouthConShr_r + WorkerConShr_r + PensionerConShr_r$$

The difference between that sum of shares and its initial value provides basic information on how the changing age structure of a population across the three groupings will typically affect consumption intensity. Multiplying that by the permanent disposable household income (HHDIsIncPerm) suggests whether consumption might likely rise or fall relative to our more inertial estimate of it, and by how much. The Aging Adjustment (normally populations are becoming older, but not necessarily so the term could be negative) is added to our earlier and more basic calculation of consumption.

$$AgingAdjustment_r = (ComputedCon_r - ComputedCon_{r,t=1}) * HHDIsIncPerm_{r,h}$$

$$C_r = C_r + AgingAdjustment_r$$

This revised consumption can then be divided among the population-age categories.

$$CPREWORK_r = \frac{YouthConShr_r}{ComputedCon_r} * C_r$$

$$CWORKING_r = \frac{WorkingConShr_r}{ComputedCon_r} * C_r$$

$$CRETIRE_r = \frac{PensionerConShr_r}{ComputedCon_r} * C_r$$

A normalization process assures that the three terms sum to C.

3.3.7.4 The impact of life expectancy on consumption

The logic around changing consumption in response to age structure does not take into account life expectancy. To do that a life expectancy term (LifExpTerm) further adjusts consumption (and also gross capital formation, IGCF). That term responds to a life expectancy gap (LifExpGap) compared to the value of the gap in the first year; the gap itself compares actual life expectancy (LifExp) with a value for expected value based on GDP per capita at PPP (GDPPCP).

$$C_r = C_r - LifExpTerm_r$$

where

$$LifExpTerm_r = IGCF_r * (LifExpGap_r - LifExpGap_{r,t=1}) * 0.01$$

$$LifExpGap_r = LifExp_{r,t-1} - LifExpGDPBased_r$$

$$LifExpGDPBased_r = F(GDPPCP_r)$$

3.3.7.5 Interest rates, stocks (inventories) and consumption

The trade-off for households between consumption and savings also responds to prices, both those of financing (interest rates) and those of goods and services. In IFs, both of those prices in turn respond to levels of inventories in the production/consumption sectors (called stocks (ST) inside IFs).

The modeling of the influence of interest rates could be done with the primary target being household savings. We instead represent the impact of rates on consumption, making savings the residual. The impact of interest rates on consumption is via an interest rate multiplier term (IntrMulTermC) that rise or falls over time as a function of the difference between a smoothed interest rate term (SmoothIntr) and a very long-term or highly-smoothed interest rate term (LongTermIntr). That is, as the smoothed interest rate rises above the long-term interest rate, it depresses consumption (and will therefore raise savings). In IFs there is no monetary sector and interest rates are real rates, not nominal ones.

Although we could similarly use the real prices of goods and services to affect consumption, the recursive structure of the model means that prices are computed in the supply/demand balance and not available at the time of this adjustment to consumption (the same lag issue affects interest rates, but the lagged values change slowly and using them is not a problem). To bypass the lag, the implicit price effect on consumption is introduced directly via an additive term reflecting half of excess stocks; the reason for passing through an adjustment of that magnitude is that household consumption is generally well above have of the total economy. Excess stocks are the difference between actual and desired stocks, which are determined by a stock base (STBase) linked to production and consumption levels and a desired stock level (dstl) as a portion of the stock base.

$$C_r = C_r * IntrMulTermC_r + ExcessStocks_r / 2$$

where

$$IntrMulTermC_r = IntrMulTermC_{r,t-1} * F(SmoothIntr_r - LongTermIntr_r, ADJSTR)$$

$$ExcessStocks_r = \sum^s (ST_{r,s} - STBase_{r,s} * dstl)$$

3.3.7.6 Savings rates and consumption

Were IFs to make savings rates responsive to interest rates and prices, rather than making consumption responsive to them, the formulation would most likely identify some long-term target value of savings rates and allow the actual value to rise above or fall below that based on driving values such as interest rates. Since IFs works primarily on consumption and makes savings the residual, it would be possible that savings rates are pushed unreasonably high or low. Moreover, it is possible that foreign savings swings can dramatically influence savings totals in a fashion that would completely crowd out or greatly expand domestic savings were there no other forces working at equilibrating domestic savings.)

To avoid this we make an additional adjustment to consumption using a multiplier on it ($MulCon$) that is responsive to the level of household and firm savings ($HHSav$ and $FIRMSav$) as a portion of relative price adjusted potential GDP ($GDPOTRPA$). That current household and firm savings rate ($HHFirmSavR$) is compared with a targeted one ($HHFirmSavTarget$) and the multiplier is computed in our standard IFs PID adjuster function ($Adjstr$). The target rate also uses one of our standard IFs functions, namely $ConvergeOverTime$, which causes it to gradually move from the initial rate to 30 percent over 80 years.

$$C_r = C_r * MulCon_r$$

where

$$MulCon_r = Adjstr(HHFirmSavR_r - HHFirmSavTarget_r)$$

$$HHFirmSavR_r = \frac{\sum^H HHSav_{r,h} + FIRMSav_r}{GDPOTRPA_r}$$

$$HHFirmSavTarget_r = ConvergeOverTime(HHFirmSavR_{r,t=1}, 30, 80)$$

3.3.7.7 Investment and consumption

The introduction of a scenario to change investment/gross capital formation (IGCF) can be done in IFs using the multiplicative parameter on investment ($invm$). That parameter varies around its base value of 1. Rapid scenario changes can complicate model adjustment. Therefore, an adjustment term ($IAdj$) is calculated based on the difference between the parameter and 1.0, and that adjustment factor is subtracted from consumption.

$$C_r = C_r - IAdj_r$$

where

$$IAdj_r = IGCF_r * (invm_r - 1)$$

Of course, the same amount needs to be added to investment.

$$I_r = I_r + IAdj_r$$

3.3.7.8 Consumption by household type and update of consumption propensity for future years

Having modified total household consumption (C) the preliminary household consumptions (ConsumHHPrel) of the different household types will no longer sum to it, so they are normalized to do so.

$$ConsumHH_{r,h} = \frac{ConsumHHPrel_{r,h}}{\sum^H ConsumHHPrel_{r,h}} * C_r$$

The rates of consumption relative to disposable income (CRA) can then be re-computed and saved for the next year for use in the preliminary calculations of consumption prior to the various adjustments of C.

$$CRA_{r,h} = \frac{ConsumHH_{r,h}}{HHDispInc_{r,h}}$$

The split of household consumption between unskilled and skilled households is maintained as a constant in this process. That is not ideal, but we have developed no alternative formulation at this point. But is therefore also important to explain the initialization of those rates of consumption by household type. In the first time step, those rates are calculated taking into account the differential propensity of populations at different levels of income to consume and save.

The average propensity to consume in the initial year (AveConsumR) can be calculated from data.

$$AveConsumR_{r,t=1} = \frac{C_{r,t=1}}{\sum^H HHInc_{r,h,t=1}}$$

If both types of households consumed an equivalent share of income, they would consume at the average rate. Almost certainly, however, lower income, unskilled households have a higher average propensity to consume than do skilled households. In the absence of data-based knowledge about that, IFs currently uses a stylistic, and flexibly changeable function to represent the differential consumption propensity at different income levels, assumed to decrease as income (proxied by GDP/capita at PPP) increases.

The consumption differential from the function adjusts the average rate for unskilled households and allows computation of actual consumption (ConsumHH) of the unskilled and a residual calculation of consumption for the skilled (subject to tests for positive sign and reasonable size, not shown below).

The differential propensity to consume is an important feature for scenario analysis, because transfers across household types (and resultant changes in relative disposable income) are one of the key policy levers available to governments.

$$AveConsumptionR_{r,h=1,t=1} = AveConsumR_{r,t=1} + ConsumDiffR_{r,t=1}$$

$$ConsumHH_{r,h=1,t=1} = HHINC_{r,h=1,t=1} * AveConsumptionR_{r,h=1,t=1}$$

$$ConsumHH_{r,h=1,t=2} = C_{r,t=1} - ConsumHH_{r,h=1,t=1}$$

where

$$ConsumDiffR_{r,t=1} = TablFunc(GDPPCP_{r,t=1})$$

This allows then that initial year estimate of consumption propensity that will vary by household type with that variation carrying over across time, even as the rate of consumption for each household type rise and fall together as the consumption share of household income rises and falls.

$$CRA_{r,h,t=1} = \frac{CONSUMHH_{r,h,t=1}}{HHDISPINC_{r,h,t=1}}$$

3.3.7.9 Household savings

Savings terms are calculated as residuals for all agent classes, and are determined after all income and expenditure/transfer calculations. In the case of households they are income, minus consumption and net transfers to government.

$$HHSav_{r,h} = HHInc_{r,h} - CONSUMHH_{r,h} - HHTax_{r,h} - HHGovSS_{r,h}$$

IFs has begun to track assets of households in total over time as an element of looking at retirement financing, but not yet begun to do that by household type so as to add that information to the behavioral representation of household-specific savings. That might be useful in the future.

3.3.8 **Household Consumption by Sector**

Consumption by sector of origin is complicated by the fact that households significantly change their pattern of consumption as their incomes increase. To capture that dynamics, IFs relies upon a linear expenditure system (LES)—for a description of the LES approach see Taylor (1979: 219-223). Consumption in each sector of origin and household type (CSHH) in the LES depends upon a minimum level of expenditure in each sector (MINEX) and the marginal budget share (MARBD) of each sector in total consumption (C) above the sum of all minimum expenditures—the floor (FLOOR). IFs computes the parameters for this consumption function during the first time cycle of the model. The constant of 0.5 in the equation for MINEX represents a Frisch parameter (the marginal utility of income with respect to income) of -2.0. Parameters and calculations are specific to household type.

$$CSHH_{r,s,h} = MINEX_{r,s,h,t=1} + MARBD_{r,s,h,t=1} * (ConsumHH_{r,h} - FLOOR_{r,h,t=1})$$

where

$$CSHH_{r,s,h,t=1} = ConsumHH_{r,h,t=1} * CS_{r,s,t=1} / C_{r,t=1}$$

$$MARBD_{r,s,h,t=1} = \frac{engel_s * \frac{CSHH_{r,s,h,t=1}}{CONSUMHH_{r,h,t=1}}}{\sum^S \left(engel_s * \frac{CSHH_{r,s,t=1}}{CONSUMHH_{r,h,t=1}} \right)}$$

$$MINEX_{r,s,h,t=1} = CONSUMHH_{r,t=1} * \left(\frac{CSHH_{r,s,t=1}}{CONSUMHH_{r,h,t=1}} - 0.5 * MARBD_{r,s,h,t=1} \right)$$

$$FLOOR_{r,h,t=1} = \sum^S MINEX_{r,s,h,t=1}$$

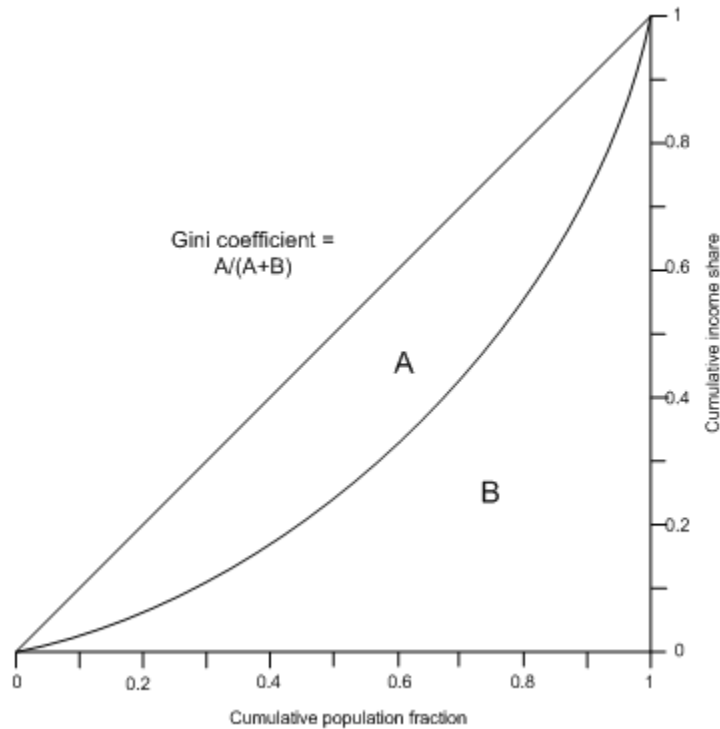
In future years, given consumption by household type and sector, total consumption by sector across households is possible to compute, and that value is necessary for the goods and services market module.

$$CS_{r,s} = \sum^H CSHH_{r,s,h}$$

3.3.9 Domestic Distribution of Household Income

3.3.9.1 Conceptual foundations of distribution in IFs: The Lorenz Curve and Gini Coefficient

The **Lorenz curve** is the most widely used method for representing inequality in earnings, income, or wealth (see the figure below). It portrays the *cumulative* share of income (or any other quantity distributed across a population) held by increasingly well-to-do *cumulative* shares of population. The more equally distributed a factor is, the closer the Lorenz curve will be to the hypotenuse of the right-triangle, sometimes called the line of equality. The **Gini coefficient** is the “area of inequality” immediately below the hypotenuse (A) divided by the area of the triangle (A+B); thus larger Gini coefficients indicate greater inequality.



The Lorenz curve is “non-parametric” in the sense that it is an empirical distribution that is an accurate representation of survey data on income or consumption for a society. While the Lorenz curve is useful conceptually, to capture the dynamically evolving distribution of income or consumption it is more convenient to have an analytic or “parametric” representation of the distribution. Moreover, we would like a representation from which we can conveniently compute specific deciles or quintiles (thereby reconstructing the Lorenz curve) and also compute key poverty measures like the headcount. The most widely used parametric representation is the lognormal density. We will not describe that function here. Because we relate the Gini coefficient to a lognormal function in our forecasting of poverty rates, please see the documentation of poverty forecasting for that discussion. Here we will focus only on the forecasting of the Lorenz curve in IFs and the calculation of the Gini coefficient from that forecasting.

The challenge for IFs forecasting of distribution becomes the forecasting of the Lorenz curve itself; calculation of Gini is simple once we have that curve at any time point for any population. Forecasting the Lorenz curve necessarily involves the forecasting of the differential performance of segments within the population. Based on historic data, the income means of different deciles could be extrapolated or in some other way forecast, but such extrapolation would be unrelated to the rest of the IFs model (including specific interventions that might change those means) and would add nothing much simply to extrapolating changes in Gini itself.

Ideally, the forecasting of segments of population from poorest to richest across the Lorenz curve should be tied to an elaboration of household types as is done with a Social Accounting Matrix (SAMs). SAMs can distinguish multiple types of urban and rural households and their changing demographic sizes (as a result, for instance, of rural-to-

urban migration) and income structures (as a result of structural change in the economy, of changing patterns of government transfers, and many other factors). For the purposes of studying longer-term change in global poverty levels in IFs the SAM ideal is tarnished by two realities: (1) in spite of efforts via the UN's System of National Accounts, there is no standard household classification system for SAMs; they vary widely and are generally used in single-country analysis, not global forecasting; and (2) most models built around SAMs are used for rather short-term analysis and even more commonly for comparative static analysis (for example, comparison of income patterns in a society not open to agricultural imports to one that is, without much or any consideration of the dynamic path from one to the other).

Nonetheless, the basic rooting of forecasts of a Lorenz curve and therefore of Gini in a SAM retains the tremendous advantage of tying those forecasts to changes in clearly policy-relevant interventions. And fortunately, the Global Trade and Analysis Project (GTAP) has collected key information, such as share taken of value added, about two classes of households, those based respectively on skilled and unskilled labor, across the large number of countries/ regions and economic sectors that the project covers (see <https://www.gtap.agecon.purdue.edu/default.asp>). IFs has drawn heavily on the GTAP data, updating regularly to the most recent version, in its own economic model structure and built the IFs SAM on the basis of the two household categories. The GTAP database has made it possible to develop forecasts of income by type of household as economies and therefore value added shares shift from agriculture to manufacturing and to services. (See the documentation topic on Household and Firm Income.)

Although it is far from ideal to have only two categories of population (unskilled and skilled), it is possible to represent a Lorenz Curve showing what portion of income the unskilled (poorer) portion of the population obtains, specifying one point on the Lorenz curve that will fall below the diagonal (the line of equality). The unskilled and skilled portions together will receive all of the income. Although simple this Lorenz curve (actually two line segments), allows calculation of Gini.

For the purposes of forecasting, we need only forecast two things: the changing shares of income of the two groups and the changing sizes of the two groups. The GTAP database facilitates the forecasting of income shares because it indicates income patterns for the two population groupings by economic sector and, across all of the countries in GTAP, helps us abstract changing patterns of those sector shares with change in GDP per capita at PPP. The documentation of household income by household type describes how that is done in IFs.

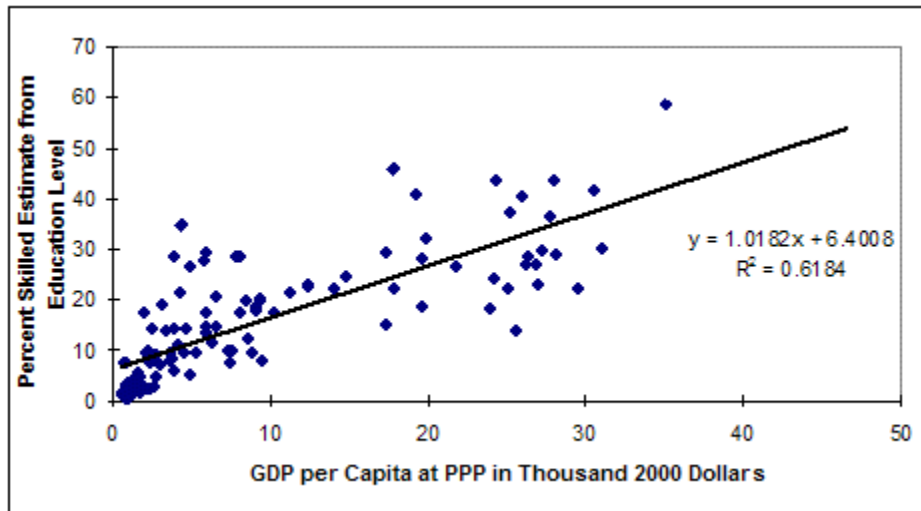
This leaves one further challenge for forecasting, the determination of the split of the population between unskilled and skilled households. Unfortunately, the GTAP dataset does not provide numbers on labor force size, only on sector income shares. And OECD data on labor force size by classification such as professional and administrative, which could be used to estimate numbers of skilled versus unskilled workers, exist only for well-to-do countries. Moreover, we need not only initial conditions, but forecasting formulations. Although we would expect that the portion of the population that is skilled would grow with the educational level of the population, the relationship is complex, in

part because (1) more advanced economies may set a higher bar for determining what a skilled worker is and (2) more highly educated populations may find that there is credential inflation that effectively demotes some with education to a less skilled category.

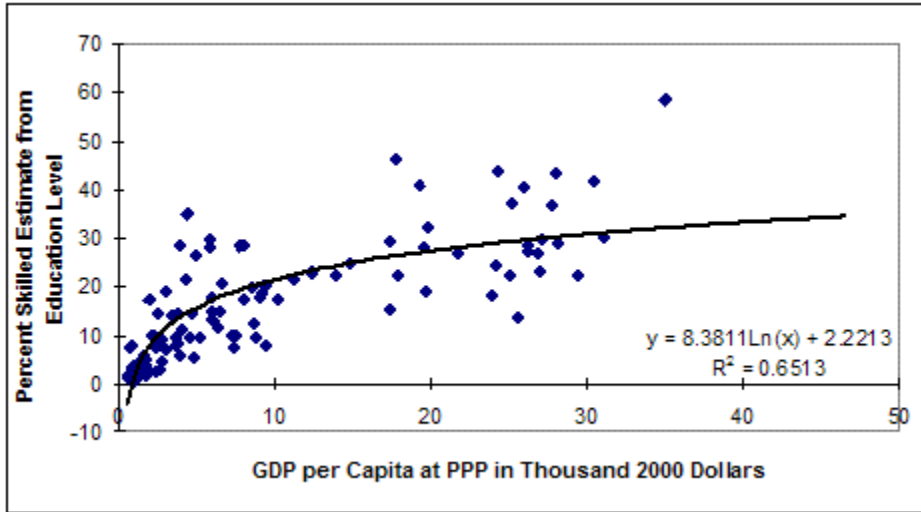
Therefore formulations and algorithms for both the initialization and forecasting of skilled and unskilled numbers are necessary. IFs has used two different processes to help generate rough estimates of the numbers, a legacy procedure that we document because some variation of it may still have value in the future, and a second, currently used process.

3.3.9.2 An earlier, legacy process to compute domestic Gini

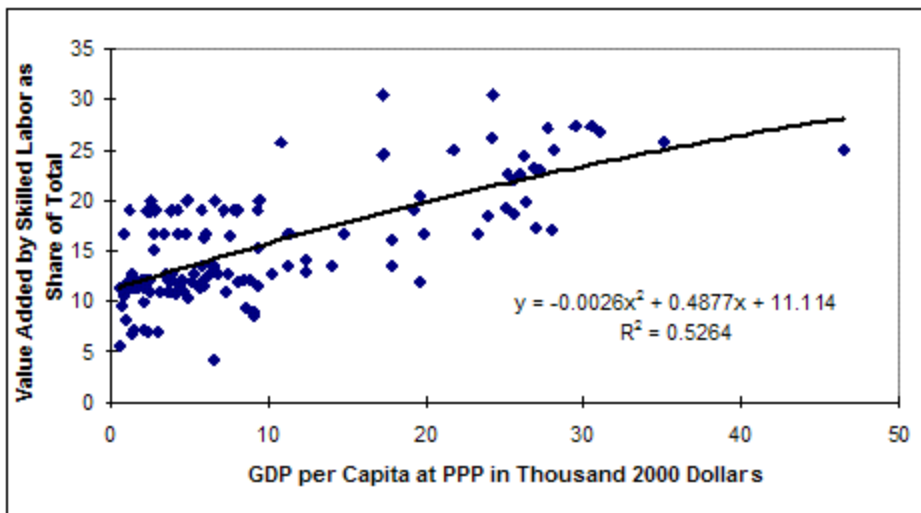
The earlier process involved looking at the formal educational attainment levels of the adult population. The cross-sectional graph below shows the portion of the population 25 years and older with tertiary education plus ½ of that with secondary education. That is a very arbitrary and crude definition of skilled labor, but is a starting point. The value of tying the size of skilled and unskilled labor forces to education completion is that the education module of IFs forecasts sizes of population with various educational attainment levels and those can be driven by agent-class interventions. And it is the change in share of skilled labor that is most important, not the absolute numbers—the Gini coefficient data can be used to adjust the initial conditions of income distribution and the change in labor force sizes can help us forecast the change in Gini.

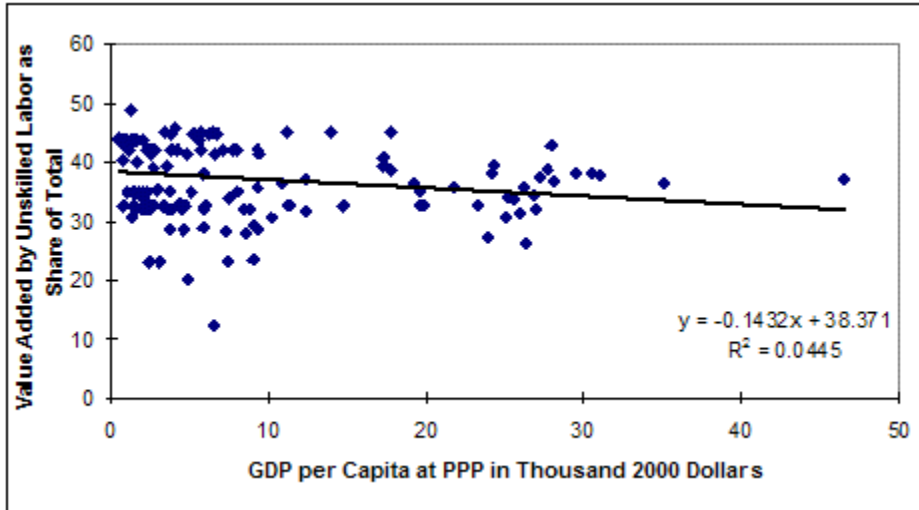


There are some obvious problems with the graph above. One of the most significant is that better fits are achieved by other functional forms, including the logarithmic one below. Unfortunately, those forms tend to suggest that the poorest countries have absolutely no skilled labor, clearly an exaggeration.



How do we know that an absence of skilled labor in the poorest countries is an exaggeration? In addition to common sense, it can be seen from the GTAP data on the income share of skilled labor in the graph below. That shows that they obtain a substantial share of value added even at the lowest levels of GDP per capita. The two cross-sectional graphs below show the shares of value added accruing to skilled and unskilled labor, respectively.





3.3.9.3 The current process in IFs to compute domestic and global Gini coefficients

An earlier procedure for computing Gini linking skilled labor to those with a tertiary education and half of those with a secondary education was formerly used in IFs. But it had some instability and unanticipated consequences, in part because the reality is there is credential inflation; as countries become richer, fewer and fewer of those with tertiary education, much less secondary, are considered adequately skill. Therefore the model has been switched to one that looks at the average years of education held by adults 15 years of age and over (EDYRSAG15).

The process begins by using a function estimated from the relationship between GDPPCP and GTAP data on the share of value added accruing to skilled workers to estimate the share of the labor force that is skilled (PerSkilled); that will be wrong, but remember that we do not need the exact share, what we need is the change over time. And the model uses that same function over time to estimate the expected share of skilled labor, all else equal, over time. But all is not equal and it is unequal to the extent that education levels rise higher than those that we would expect in the same development process. Hence there is an adjustment term in the calculation (LabSupSkillAdj) that looks at the difference between the actual years of education in the population and the expected years (YearsEdExp), again computed by an analytical function with GDP per capita at PPP. A fraction of the proportional difference of actual with expected is added to the percentage skilled.

$$PerSkilled_r = PerSkilledExp_r * (1 + LabSupSkillAdj_r)$$

where

$$PerSkilledExp_r = AnalFun(GDPCPP)$$

$$LabSupSkillAdj_r = \frac{EDYRSAG15_r - YearsEdExp_r}{YearsEdExp_r} * 0.3$$

$$YearsEdExp_r = AnalFun(GDPCPP)$$

Given the percentage skilled it is simple to calculate the number skilled (LABSUP) and the residual percentage of the labor force that is unskilled.

$$LABSUP_{r,h=2} = LAB_r * PerSkilled_r / 100$$

$$LABSUP_{r,h=1} = LAB_r - LABSUP_{r,h=2}$$

Levels of skill in the labor force are used to estimate household populations with associated skills (recognizing that household sizes in different skill categories may actually be unequal).

$$HHPOP_{r,h} = POP_r * LABSUP_{r,h} / LAB_r$$

Given the income shares accruing to skilled and unskilled shares of the population and the sizes of those portions of the population, the domestic Gini index (GINIDOM) is computed from the simple Lorenz curve that those two incomes and population shares create, scaled to the empirically known initial condition. The calculation in the Lorenz curve function `GinDomCalcFunc` is modified by a scaling or shift parameter (`GINIDomRI`) so that the initial condition is preserved. A multiplier parameter (***ginidomm***) can change the calculation over time.

$$GINIDOM_r = (GiniDomCalcFunc(HHINC_{r,h} HHPOP_{r,h}) + GINIDomRI_{r,t=1}) * \mathbf{ginidomm}_r$$

where

$$GINIDomRI_{r,t=1} = GINIDOM_{r,t=1} - GiniDomCalcFunc(HHINC_{r,h,t=1} HHPOP_{r,h,t=1})$$

There is, however, an alternative and completely exogenous way of calculating Gini, if the parameter ***ginidomr*** has a value greater than 0. In that case, the value of Gini is simply the initial condition changed over time by ***ginidomr*** (again modified, if desired, by ***ginidomm***).

$$GINIDOM_r = GINIDOM_{r,t=1} * \mathbf{ginidomr}_r * \mathbf{ginidomm}_r \text{ if } \mathbf{ginidomr}_r > 0$$

IFs also calculates two variations of a global Gini index to supplement the country-specific one. The first is a Gini coefficient (GINI) that describes the distribution of income across all countries, using country incomes as the unit of analysis.

$$GINI = GiniCalc(HHINC_{r,h}HHPOP_{r,h})$$

That measure does not take into account income distribution within countries. The second global Gini measure (GINIFULL) does take into account income distributions both within and across countries (subject, of course, to the limitation that IFs represents domestic income distribution only as a function of two income categories). That makes it essentially a measure of global income distribution at the individual human level, rather than the country level.

$$GINIFULL = GiniFullCalc(HHINC_{r,h}HHPOP_{r,h})$$

Finally, the user interface of IFs uses the same Lorenz-curve approach to allow the user to calculate a specialized-display GINI for any variable that can be represented across all countries/regions of the model.

3.3.10 *Firm Savings*

The agent class that is least developed within IFs as a class with behavioral relationships, and that instead relies on aggregated relationships is the firm (as discussed around the production function). Even if the basic behavior of the firm is simplified, its accounts must still balance. Firm savings is equal to firm income minus net transfers to households and government. In addition, there are positive net transfers to the rest of the world in the form of outward FDI flows (XFDIFOUT), portfolio flows (XPORTFOUT), and subscriptions to the World Bank (XWBSUBF) by net contributors.

$$FirmSav_r = FirmInc_r - \sum^H HHDivInt_{r,h} - FirmTax_r - IndirectTax_r - FirmGovSS_r - XFDIFOUT_r - XPORTFOUT_r - XWBSUBF_r$$

3.3.11 *Domestic Closure: Total Savings, Investment, and Interest Rates*

Given information about the financial accounting within government, households, and firms, it is possible to turn to total domestic economy balances and also to consider the linkages with the global financial system. We want to look in turn at savings, investment and interest rates, then begin to identify some important elements of international connections.

3.3.11.1 Foreign and total savings

Other sections of this documentation explain the computations of government, household, and firm savings. The only element missing for the calculation of total savings is foreign savings (savings from abroad). Foreign savings is basically the flipside of the current account—remember that a deficit in the current account must be offset by an inflow of

foreign savings. Hence everything that improves the current account, such as more exports, reduces the net inflow of foreign savings. In IFs foreign savings have terms from trade and aid, but also from the balances of government and firms with the outside world. For government those terms involves credits and flows to and from the IMF (XIMFCRFOUT and XIMFCRFIN) and the World Bank (XWBLNFOUT and XWBLNFIN) For firms those terms involve foreign direct investment (XFDIFOUT and XFDIFIN) and portfolio flows (XPORTFOUT and XPORTFIN) as well as subscriptions to the World Bank institutions (XWBSUBF) and worker remittances (XWORKREMIT).

$$ForSav_r = GovForSavBal_r + FirmForSavBal_r - X_r + M_r - AID_r$$

where

$$GovForSavBal_r = XIMFCRFOUT_r - XIMFCRFIN_r + XWBLNFOUT_r - XWBLNFIN_r$$

$$FirmForSavBal_r = XFDIFOUT_r - XFDIFIN_r + XPORTFOUT_r - XPORTFIN_r + XWBSUBF_r - XWORKREMIT_r$$

The total savings in each country will be the sum of the individual terms. By definition it will equal investment (I), defined as capital formation (IGCF) plus inventory stock (ST) changes. Treatment of physical balances over time, not elaborated in this topic, assure that equality. Looking at the two variables side by side is a good test of the functioning of the SAM.

$$Savings_r = \sum^H HHSav_{r,h} + FirmSav_r + GovBal_r + ForSav_r$$

In other sections of this documentation we have described the computation of relative-price adjusted variables such as exports (XRPA), imports (MRPA), and GDP, potential and actual (GDPRPA and GDPPOTRPA). In addition we have relative price adjusted versions of foreign savings (FORSAVRPA) and total savings (SAVINGSRPA).

$$FORSAVRPA_r = XRPA_r - MRPA_r - AID_r$$

$$SAVINGSRPA_r = \sum^H HHSav_{r,h} + FirmSav_r + GovBal_r + FORSAVRPA_r$$

These relative-price adjusted savings variables are used for information only and do not affect financial balances or other variables. The relative-price adjusted trade, however, is taken to the trade balance (TRADEBAL) in the economic model and on to the current account balance (CURACT) in the international financial accounting and equilibration. There it will affect international debt calculations (XDEBTRPA).

3.3.11.2 Gross capital formation and investment

The equation below indicates that the core of the gross capital formation (IGCF) calculation each year is potential GDP (GDPPOT) times an investment rate (IRA) from the initial year. The investment rate core determination also involves input from foreign direct investment, increasing or decreasing if the FDI ratio to potential, relative price adjusted GDP changes.

$$IGCF_r = GDPPOT_r * (IRA_{r,t=1} + FDITerm_r)$$

where

$$FDITerm_r = \frac{XFDIFIN_{r,t-1} - XFDIFOUT_r * xfdioutf}{GDPPOTRPA_r} - \frac{XFDIFIN_{r,t=1} - XFDIFOUT_{r,t=1} * xfdioutf}{GDPPOTRPA_{r,t=1}}$$

There are multiple factors that modify that core. (See the similarities in the calculation of household consumption as having a core in permanent income, with a number of factors modifying it.) In general, those modifications help investment react to changes in savings patterns and they help maintain the equilibrium of the domestic financial system. The first is an interest rate multiplier (IRAMul) that, like the one on consumption, drives down investment when interest rates rise. In fact, the multiplier responds to smoothed and long-term interest rates exactly as consumption responds, although the parameters in the PID adjustment function make the adjustment a little less responsive. The second is one that ties investment to a smoothed version of the SAVINGS variable. As savings rise, normally investment does also, and IFs uses a moving average to smooth the impact of changes of savings rates in this relationship. The third is a direct response to excess stocks or inventories. Again this is parallel to the adjustment of household consumption to stocks, except that 1/2 of the excess stocks are passed to household consumption and only 1/6 of them are assigned to gross capital formation.

$$IGCF_r = IGCF_r * IntrMulTermI_r * SmoothSavingsTerm_t + ExcessStocks_r / 6$$

where

$$IntrMulTermI_r = IntrMulTermI_{r,t-1} * F(SmoothIntr_r - LongTermIntr_r, ADJSTR)$$

$$SmoothSavingsTerm_t = \frac{SmoothSavingsR_r}{SmoothSavingsR_{r,t=1}}$$

$$SmoothSavingsR_r = MovingAverage\left(\frac{SAVINGS_r}{GDPPOT_r}, .8\right)$$

$$ExcessStocks_r = \sum^s (ST_{r,s} - STBase_{r,s} * dstl)$$

Additionally, there are two adjustments that reflect the changing life expectancy and/or savings patterns of the society. The first reacts to changes over time in the variation of the actual life expectancy from the life expectancy expected given the GDP per capita at PPP of a society. If that gap increases, a small portion of it increases investment rates (as it would savings rates). The second again involves life expectancy, but the purpose of it is to reflect the changing age structure of a society and the consumption/savings-investment patterns of societies with different age structures. Finally, an exogenous multiplier (*invm*) allows the user to push capital formation rates up or down.

$$IGCF_r = IGCF_r * invm_r + IGCF_r * (LifExpGap_r - LifExpGap_{r,t=1}) * .01 + SavingsAgeAdj_r$$

where

$$LifExpGap_r = LIFEXP_r - LifExpComp_r$$

$$LifExpComp_r = F(GDPPCP_r)$$

$$SavingsAgeAdj_r = F(LIFEXP_r, Age Structure, Savings Structure)$$

Gross capital formation, although highly correlated with investment (I) is not equal to it. Investment also contains changes in stocks. It is that addition of delta stocks that brings the entire representation of the goods and services market in proper relationship with the financial representations of flows among agent classes.

$$I_r = IGCF_r + \sum^s (ST_{r,s,t} - ST_{r,s,t-1})$$

Having total investment, it is also possible to compute investment by origin sector (INVS). In IFs, that spread is constant over time in the configuration of the first model year. See documentation on the goods and services market for information on investment by destination sector, IDS.

$$INVS_{r,s} = I_r * \frac{INVS_{r,s,t=1}}{I_{r,t=1}}$$

3.3.11.3 Interest rates

The SAM structure in IFs is really a combination of an accounting system and an equilibrating system. Just as the equilibrating mechanisms discussed above are important, it is critical to make sure that the accounting balances are maintained. One good measure of ultimate balance is the required, conceptual equality of total savings and investment. Their equality, in fact, is also reassurance that the goods and service market elements of the model are fully integrated with the broader financial ones.

Just as changes in real relative prices by sector help the goods and services market chase equilibrium across time (using build-up or run-down of stocks or inventories to change

prices and send signals to supply and demand sides of the model), changes in real interest rates (INTR) send signals to consumption and investment in order to allow the system to chase equilibrium between savings and investment/consumption over time. IFs does not maintain a monetary sector that would introduce nominal interest rates (or nominal prices).

IFs slows down change in interest rates by making them responsive to the average of a term that is directly responsive to inventory levels and the interest rate in the preceding year. It further anchors interest rates in the responsive term to the real interest rate of the initial model year (because this is set at a stylized rate of about 3 percent rather than to actual nominal rates in the initial year, the anchor does not need to converge to a long-term target level). The responsiveness comes from a ratio of a sum across all sectors of the desired stocks (a stock base, STBase, times a parameter of desired stock rates, *dstl*) to a sum across sectors of the actual stocks (ST). Because the denominator in the ratio could become inappropriately small, it is bound so as to be at least 30 percent of desired stocks. See documentation on changes in relative prices for more information on the stock base and the general approach. Interest rates rise as stocks fall relative to desired levels and fall as stocks rise above those desired levels.

$$INTR_r = (INTR_{r,t=1} * \frac{\sum^S STBase_{r,s} * dstl}{AMAX(\sum^S STBase_{r,s} * dstl * 0.3, \sum^S ST_{r,s})} + INTR_{r,t-1})/2$$

where

$$STBASE_{r,s} = ZSP_{r,s} + MS_{r,s,t=1} * \frac{GDPPOT_r}{GDP_{r,t=1}}$$

Even with the smoothing of the above equation, a model of long-term behavior like IFs benefits from some further smoothing of the interest rates before they feed back to the consumption and investment patterns. Two smoothed terms (a very long-term interest rate, LongTermIntr, and a smoothed rate, SmoothedIntr) moderate the impacts of real interest rates and allow somewhat more variation in them.

$$LongTermIntr_r = LongTermIntr_r * 0.95 + INTR_r * 0.05$$

$$SmoothIntr_r = SmoothIntr_r * 0.7 + INTR_r * 0.3$$

3.3.12 Current and Capital Account Elements

The current account includes trade in goods and services; income on assets (such as rents, interest, profits and dividends); and current transfer payments (including foreign

aid, government subscriptions to international organizations, pension payments to workers living abroad, and foreign worker remittances to their home countries). The larger of these elements have representation in IFs.

The capital account involves changes in ownership of assets, such as includes foreign direct investment (FDI), portfolio flows, IMF credits and World Bank loans, net bank lending, other net flows, and change in reserves. Note that these capital account related flows have associated stocks (assets) of importance in understanding their own long-term behavior and that of the larger financial system including the current account.

With respect to the elements of the current and capital accounts, another part of this documentation (part of the representation of goods and services and their balances) has explained trade. This section will discuss foreign aid and worker remittances, two additional elements of the current account, then move to FDI and portfolio flows, two critical elements of the capital account. It will then move to flows from and to international financial institutions, which are also placed in the capital account. A subsequent section on overall equilibration will also address changes in reserves, which are strictly speaking also part of the broader capital account.

3.3.12.1 Foreign Aid

IFs uses a "pool" approach to aid (AID) rather than indicating bilateral flows from particular donors to particular recipients. That is, all aid from all donors flows into the pool and then all recipients draw proportions of the pool.

IFs uses the aid value parameter (*aiddon*) to calculate the aid (AID) from donors and *aidrec* to calculate the targeted aid to recipients. The pool of aid donations determines the actual total level of interstate aid flows, however, and is allocated among potential recipients according to the proportions targeted for each.

$$AID_r = \frac{GDP * (aidrec_r - aiddon_r)}{100}$$

Aid outflows are negative and the total aid pool given (AIDP) is the sum of the negative flows, while the total desired aid of recipients (AIDR) is the sum of positive flows.

$$AIDP = \sum^R -AID_r \quad \text{if } AID_r < 1$$

$$AIDR = \sum^R AID_r \quad \text{if } AID_r > 1$$

A re-computation of aid for recipients distributes the aid pool across their demands.

$$AID = AIDP * \frac{AID_r}{AIDR} \quad \text{if } AID_r > 1$$

3.3.12.2 Worker remittances

Worker remittances are tied closely to the migration formulation of the IFs demographic model (see the documentation of that model). It is important to know the number of foreign workers from each country and into each country (POPFOREIGN), as well as the pattern of remittances per worker, as a basis for computing remittance levels. The IFs migration module once relied heavily on UN data for past migration levels and also tapped their near-term forecasts. The IFs preprocessor and model now rely primarily on migration data from the International Institute of Applied Systems Analysis (IIASA). The preprocessor has fairly extensive algorithms to reconcile that migration data with remittance data, and migration and remittance data from the World Development Indicators with foreign population data (from the same source).

The data do not, however, exist for global representation of dyadic remittance flows that is, of flows from country of origin to country of destination. IFs therefore uses a pooled approach, representing aggregate flows out of countries and aggregate flows into countries, requiring that they be balanced, but not attempting to trace dyadic patterns. The pooled approach is also used by IFs in trade (where dyadic representation would be possible), foreign aid (where it would be difficult but not completely impossible to obtain dyadic data), and other financial flows (where again dyadic data are seldom available).

The specific formulation computes a global average remittance rate per worker (WWorkRemitRate) and a host-country specific ratio of remittance rate to the global one (WorkRemitRate) in the first year. In subsequent years, those rates are applied to the number of foreign workers, but adjusted by the ratio of current GDP per capita to initial GDP per capita. The outflows are assigned to inflow countries so as to maintain a global balance.

$$XWORKREMIT_r = -POPFOREIGN_r * WWorkRemitRate_{t=1} * \\ WorkRemitRateI_{r,t=1} * \frac{GDPPC_r}{GDPPC_{r,t=1}}$$

where

$$WWorkRemitRate_{t=1} = \frac{\sum^R -XWORKREMIT_{r,t=1}}{\sum^R POPFOREIGN_{r,t=1}}$$

$$WorkRemitRateI_{r,t=1} = \frac{\frac{Abs(XWORKREMIT_{r,t=1})}{Abs(POPFOREIGN_{r,t=1})}}{WWorkRemitRate_{t=1}}$$

Initial data on worker remittances, added to the IFs database as a percentage of GDP, come from the World Bank's WDI, as do data on foreign population as a percentage of the total. The original sources were Global Development Finance and the OECD,

respectively. At one time the worker remittance series only included data by country of receipt, forcing us to estimate remittances by countries of origin in the preprocessor on the basis of the foreign worker series. Now remittance flows in and out are reported, but there are still balancing issues.

3.3.12.3 Foreign Direct Investment

Firms are primarily in charge on both ends of foreign direct investment. In general, of course, the pattern is likely to be that firms direct FDI from relatively capital-rich countries to relatively capital-poor ones. Estimation reinforces that presumption by showing the patterns found in the IFs database (using FDI flow data from WDI). A less steeply sloped line is the relationship between GDP per capita at PPP and the stocks of FDI inflows as a ratio to GDP. A more steeply sloped line is the relationship between GDP per capita at PPP and the stocks of FDI outflows as a ratio to GDP. Both lines are upward sloping and, in fact, countries are simultaneously larger sources and targets of investment, even relative to GDP, as they develop. Yet, roughly speaking, countries are net recipients until GDP per capita is somewhat above \$20,000 and net sources thereafter.

IFs recognizes that these patterns will not be universal. Thus the algorithm that determines stocks of investment outflows is one that builds in the historic pattern of FDI outflows, but that assumes convergence over long periods of time, such as a century, towards the generic pattern. The same is true for recipients and stocks of inflows. It would probably be reasonable to posit that both lines would shift to the right over time as the average per capita levels of global GDP increase. We have not built that presumption into the model at this time.

In addition to the relative behavior of firms in states across the system, another behavioral issue is the overall pattern of increase or decrease in FDI flows relative to the size of the global economy. Over the last several decades FDI has grown steadily as a portion of the global capital stock and global economy. Economic historians are, however, quick to point out that the turn of the 20th century was a period of enhanced globalization of capital and that those flows then retreated for most of the 20th century before advancing again. And this century has already seen some at least temporary retreats. Thus the base case presumption built into IFs, based roughly on patterns of the late 1990s, is of growth in FDI flows at a rate that exceeds economic growth but that convergences towards GDP growth by 2010.

The representation of FDI in IFs builds from the concepts and general theory around FDI stocks that estimation and the preceding discussion suggest. In each time cycle the model computes stocks of both inflows and outflows. As a first step, the expected stocks of inflows (*EXFDISTOCK*) are a fraction of GDP (*FDIRatioIn*), where the ratio gradually converges from the initial condition for each country (*XFDIGDPIIn*) to the pattern of global ratios expected as a function of GDP (*ExpFDIRatioIn*).

$$EXFDISTOCK_r = GDP_r * FDIRatioIn$$

where

$ExpFDIRatioIn = AnalFunc(GDPPCP)$

$$XFDIGDPIn_{r,t=1} = \frac{XFDISTOCK_{r,t=1}}{GDP_{r,t=1}}$$

$FDIRatioIn = ConvergeOverTime(XFDIGDPIn_{r,t=1}, ExpFDIRatioIn, 100)$

Exactly the same logic applies to expected stocks of outflows (EXFDISTOUT).

$$EXFDISTOUT_r = GDP_r * FDIRatioOut$$

where

$ExpFDIRatioOut = AnalFunc(GDPPCP)$

$$XFDIGDPOut_{r,t=1} = \frac{XFDISTOUT_{r,t=1}}{GDP_{r,t=1}}$$

$FDIRatioOut = ConvergeOverTime(XFDIGDPOut_{r,t=1}, ExpFDIRatioOut, 100)$

Without further modification, the above formulations would result in stocks that fully determined the annual inflows and outflows. There are, however, empirical initial conditions for such annual flows that should not be ignored, especially in the early years of forecasts; the late 1990s were, for instance, a period of high flow rates that would not be captured by the above formulations.

Thus the formulation has a second step built around flows. Based on the above-expected stocks of FDI inflows (EXFDISTOCK), actual stocks (XFDISTOCK) are computed using a rate of growth (FDIRIn). The user can further intervene in the stock specification with a multiplier (*xfdistockm*). The rate of FDI stock growth converges from an initial rate tied to empirical inflows to an exogenous rate (*xfdistockr*).

$$XFDISTOCK_r = EXFDISTOCK_r * \left(1 + \frac{FDIRIn}{100}\right) * xfdistockm_r$$

where

$$XFDISTOCKR_{r,t=1} = \frac{XFDIFIN_{r,t=1}}{XFDISTOCK_{r,t=1}} * 100$$

$FDIRIn = ConvergeOverTime(XFDISTOCKR_{r,t=1}, xfdistockr_r, 100)$

Again the logic is the same for outward flows and stocks.

$$XFDISTOUT_r = EXFDISTOUT_r * \left(1 + \frac{FDIROut}{100}\right) * xfdistockm_r$$

where

$$XFDISTOUTR_{r,t=1} = \frac{XFDIFOUT_{r,t=1}}{XFDISTOUT_{r,t=1}} * 100$$

$$FDIOut = ConvergeOverTime(XFDISTOUTR_{r,t=1}, \mathbf{xfdistockr}_r, 100)$$

There is, therefore, in the formulation a tension between the convergence of stock patterns to the analytic function for stocks as a portion of GDP and the convergence of them towards growth as specified in the exogenous parameter for rate of stock growth (*xfdistockr*). If the model user were to set the exogenous parameter for all geographical units and time points to approximately the rate of global economic growth, the tension would be resolved. Many users will, however, want to override the analytic specification by using that exogenous parameter.

There is one additional element of the formulation for FDI, which is primarily in place for the purpose of model use in scenarios. It will often be desirable for a user to be able to simply specify a rate of global growth in FDI (including potentially sharp retrenchments in FDI) and impose the resulting global FDI levels (WFDI) on the above formulations, letting the above equations determine country/regional stocks and flows within the global constraint. This formulation element uses a rate of growth (WFDIR) in the ratio of FDI to world GDP (WFDIRGDP) that is set to move from the initially calculated growth rate (WFDIGRI) of world foreign direct investment relative to GDP growth to 0.5% (now WFDIGR) over the first 20 years of the model run. The calculation takes into account the expected depreciation rate of the FDI stock as represented by the lifetime of manufacturing capital (*lks*). The user has some control over this global growth pattern in the form of a multiplier on global FDI (*wfdiwgrm*). Once WFDI is determined, it is used to normalize the country/regional calculations of stocks (XFDISTOCK and XFDISTOUT)—normalization not shown in the equations.

$$WFDI = WFDIRGDP_{t-1} * WGDP * (1 + WFDIR * \mathbf{wfdiwgrm})$$

where

$$WFDIR = ConvergeOverTime(WFDIGRI_{t=2}, WFDIGR, 20)$$

$$WFDIRGDP_{t-1} = \frac{WFDI_{t-1}}{WGDP_{t-1}}$$

$$WFDIGRI_{t=2} = \frac{WFDIFLOWS_{t=2} * (1 + WFDIGR_{t=2}) - \frac{WFDISTOCK_{t=2}}{\mathbf{lks}_{3=manuf}}}{WFDISTOCK_{t=2} - WGDPGROWTH_{t=2}}$$

$$WGDPGROWTH_{t=2} = \frac{WGDP_{t=2} - WGDP_{t=1}}{WGDP_{t=1}}$$

$$WFDIGR = 0.005$$

The changes in the ultimate values of stocks, of course, then provide gross inflows and outflows.

$$XFDIFIN_{r,t} = XFDISTOCK_{r,t} - XFDISTOCK_{r,t-1} + \frac{XFDISTOCK_{r,t-1}}{lks_{3=manuf}}$$

$$XFDIFOUT_{r,t} = XFDISTOUT_{r,t} - XFDISTOUT_{r,t-1} + \frac{XFDISTOUT_{r,t-1}}{lks_{3=manuf}}$$

Obtaining data for initialization of the stocks and flows of FDI posed some challenges. UNCTAD's annual World Investment Report had some data on stocks, but they proved difficult to obtain electronically. We therefore turned once again to the World Development Indicators, in spite of the fact that the source provides only flows. It does, however, basically provide both net inflows and net outflows (XFDIFIN and XFDIFOUT). Because of the instability in these numbers over time, the last five years of data were averaged to compute initial conditions for annual flows in the model.

Integrating the flows from 1970, on the assumption that there was relatively little FDI prior to 1970, provided estimates of stocks that seemed fairly reasonable when checked against some of the UNCTAD stock data. But because these do not necessarily yield the same totals for global inflows and outflows, the two sets of numbers were summed for their respective totals, true values were assumed to be the average, and inflows and outflows were normalized to that average. The result was initialization of FDI stocks from abroad (XFDISTOCK) and FDI stocks held abroad (XFDISTOUT).

3.3.12.4 Portfolio Flows

The algorithm behind portfolio flows, again involving firms in large part but also introducing behavior of households, is essentially the same as that for FDI. And again, model users can intervene to change either systemic growth of portfolio flows or individual country/region patterns.

With respect to portfolio investment, WDI provides flow data on both bond and equity investment. Again averages of the most recent five years were used to compute initial conditions for annual flows, and annual flow data were integrated from 1970 to obtain estimates of portfolio stocks held in developing countries (XPORTFOLIO). Unfortunately, the WDI provided only flows into developing countries and the data fail to account for the source of the flows. Therefore the asset balances were assigned by GDP weight in the preprocessor to developed countries (XPORTSTOUT).

In each time cycle the model computes stocks of both inflows and outflows. As a first step, the expected stocks of inflows (XPORTFOLIO) are a fraction of GDP (PortfolioRatio), where the ratio gradually converges from the initial condition for each country (XPORTGDPIn) to the pattern of global ratios expected as a function of GDP (ExpPortRatioComp).

$$EXPORTFOLIO_r = GDPPOT_r * PortfolioRatioIn$$

where

$$ExPortfolioRatioInComp = AnalFunc(GDPPCP)$$

$$XPORTGDPI_{r,t=1} = \frac{XPORTFOLIO_{r,t=1}}{GDP_{r,t=1}}$$

$$PortfolioRatioIn = ConvergeOverTime(XPortGDPI_{r,t=1}, ExPortfolioRatioInComp, 100)$$

Without further modification, the above formulations would result in stocks that fully determined the annual inflows. There are, however, empirical initial conditions for such annual flows that should not be ignored, especially in the early years of forecasts; the late 1990s were, for instance, a period of high flow rates that would not be captured by the above formulations.

Thus the formulation has a second step built around flows. Based on the above-expected stocks of portfolio inflows (EXPorfolio), actual stocks (XPORTFOLIO) are computed using a rate of growth (PortfolioRIn). The user can further intervene in the stock specification with a multiplier (*xportfoliom*). The rate of portfolio stock growth converges from an initial rate tied to empirical inflows to an exogenous rate (*xportr*).

$$XPORTFOLIO_{r,t} = EXPORTFOLIO_r * \left(1 + \frac{PortfolioR}{100}\right) * xportfoliom_r$$

where

$$XPORTFOLIOR_{r,t=1} = \frac{XPORTFIN_{r,t=1}}{XPORTFOLIO_{r,t=1}} * 100$$

$$PortfolioR = ConvergeOverTime(XPORTFOLIOR_{r,t=1}, xportr, 100)$$

The same logic applies to portfolio outflows, starting with a computation of expected flows.

$$EXPORTSTOUT_r = GDPPOT_r * PortfolioRatioOut$$

where

$$ExPortfolioRatioOutComp = AnalFunc(GDPPCP)$$

$$XPORTGDPOut_{r,t=1} = \frac{XPORTTOUT_{r,t=1}}{GDP_{r,t=1}}$$

PortfolioRatioOut

$$= \text{ConvergeOverTime}(XPortGDPOut_{r,t=1}, \text{ExpPortfolioRatioOutComp}, 100)$$

The expected flows out can then be modified to make them more compatible with initial data and user interventions.

$$XPORTSTOUT_r = EXPORTSTOUT_r * \left(1 + \frac{\text{PortfolioR}}{100}\right) * \mathbf{xportstoutm}_r$$

where

$$XPORTSTOUTR_{r,t=1} = \frac{XPORTFOUT_{r,t=1}}{XFPORTSTOUT_{r,t=1}}$$

$$\text{PortfolioR} = \text{ConvergeOverTime}(XPORTSTOUTR_{r,t=1}, \mathbf{xportr}, 100)$$

The world sums of stocks for portfolio inflows and outflows will not be identical. IFs computes a world total of flows (WPortfolio) as a step towards imposing it on both.

$$WPortfolio = WPortfolioGDP_{t-1} * WGDP * (1 + WPortfolioR * \mathbf{xportwgrm})$$

where

$$WPortfolioR = \text{ConvergeOverTime}(0.1, 0.005, 10)$$

$$WPortfolioRGDP_{t-1} = \frac{WPortfolio_{t-1}}{WGDP_{t-1}}$$

$$XPORTFOLIO_r = XPORTFOLIO_r \frac{WPortfolio}{WPortfolioIn}$$

$$XPORTFOLIO_r = XPORTSTOUT_r \frac{WPortfolio}{WPortfolioOut}$$

After the logic of the stock-driven formulation is played out, the final determination of flows is identical to that for FDI except that there is no physical depreciation around portfolio investment.

$$XPORTFIN_{r,t} = XPORTFOLIO_{r,t} - XPORTFOLIO_{r,t-1}$$

$$XPORTFOUT_{r,t} = XPORTSTOUT_{r,t} - XPORTSTOUT_{r,t-1}$$

Ideally foreign direct investment and portfolio investment would be responsive to equilibrating mechanisms. In reality, that is not always the case. FDI and portfolio flows can be, in fact, somewhat perverse, moving from fear or greed in opposite directions to the needed balances. At this point, no linkages of any kind have been added to these

flows from exchange rates, liquidity levels, or other elements of the equilibrating mechanisms to be discussed later.

Also, at this point there is no representation of dynamics for direct firm borrowing and lending across country borders. Instead there is a simple growth equation for net external debt. At the least, this equation should be parameterized for user intervention.

$$XFIRMDEBT_{r,t} = XFIRMDEBT_{r,t-1} * 1.03$$

3.3.12.5 International financial institution flows

The international financial institutions (IFIs) are agents of importance for countries around the world. This section explains a basic agent representation of the World Bank and a less-fully elaboration representation of the IMF.

With respect to World Bank agency, the formulations of IFs fall generally into four basic decision categories: How much to loan? To whom to loan? For what purposes to loan? At what terms to loan?

Concerning the total that the bank has to loan at any given time ($XWBLOANSTOT$), that is a simple sum of lending capacity last year plus repayments ($XWBLnRepayTot$), minus new loans ($XWBLnNewTot$), and plus new subscriptions ($XWBLnNewSubTot$), the last term modified by the loan to equity ratio maintained by the Bank ($xwblneqr$).

Repayments or the flows out of countries to the Bank ($XWBLNFOUT$) depend on total lending to the countries ($XWBLOANS$), the real interest rates charged by the Bank ($wblintr$) and the repayment rates in years ($xwblnrepr$).

The World Bank and its member countries collectively determine the rates of growth in new subscriptions ($XWBSUBF$), but we represent that with an exogenous parameter ($xwbloanr$). IFs does not attempt to maintain cohorts of loans to different countries. Nor does IFs track the administrative overhead of the Bank, which could significantly reduce lending capacity over time.

$$XWBLOANSTOT_t = XWBLOANSTOT_{t-1} + XWBLnRepayTot_t - XWBLnNewTot_t + XWBLnNewSubTot_t * xwblneqr$$

where

$$XWBLNFOUT_{r,t} = XWBLOANS_{r,t-1} * \frac{wblintr}{100} + \frac{XWBLOANS_{r,t-1}}{xwblnrepr}$$

$$XWBLnRepayTot_t = \sum^R XWBLNFOUT_{r,t}$$

$$XWBLnNewTot_t = \sum^R XWBLNFIN_{r,t}$$

$$XWBSUBF_{r,t} = XWBSUBF_{r,t-1} * \left(1 + \frac{xwbloanr_r}{100} \right)$$

$$XWBLnNewSubTot_t = \sum^R XWBSUBF_{r,t}$$

With respect to the countries to which it lends, IFs focuses primarily on determining flows as a percentage of GDP. This is different than the approach to representing FDI or portfolio investment in part because those representations were general with respect to agency. In the case of World Bank lending, there is much reason to believe that it does target flows based on characteristics of the borrowing country. Estimations show three characteristics that demonstrate correlations with annual lending by the Bank as a portion of GDP: GDP per capita (PPP), the Gini coefficient of inequality, and aid that countries receive as official development assistance (ODA) from countries. A more extensive analysis is documented elsewhere (Hughes and Hossain 2003). For instance, GDP per capita and the Human Development Index (HDI) are close substitutes in correlations with Bank lending. After extensive exploration, it was decided to add to IFs a three-variable representation of Bank decision-making concerning the countries to whom it lends.

The function detailed above provides the foundation for a three-step representation of the country-specific lending flows from the bank (XWBLNFIN). The first step is to compute a base value for flows to each country, simply by scaling up each country's initial lending receipts by the growth in total World Bank lending. Computing such a base is valuable because it preserves the bases for lending that are not captured in the predictive formulation (the three variables capture about three-fourths of the bases).

$$XWBLnFinBase_{r,t} = XWBLNFIN_{r,t=1} * \frac{XWBLOANSTOT_t}{XWBLOANSTOT_{t=1}}$$

The second step is to compute an adjusted forecast of lending, taking into account the changes in the predictive factors from initial values. The adjusting factors come into the equation as a ratio of the purely predicted lending rate in the current time period over that rate in the initial period.

$$XWBLnFinPred_{r,t} = AnalFunc(GDPPCP_r, GINIDOM_r, AIDPERGDP_r)$$

$$XWBLnFinAdj_{r,t} = \frac{XWBLnFinPred_{r,t}}{XWBLnFinPred_{r,t=1}} * XWBLnFinBase_{r,t}$$

The third and final step is to normalize the adjusted forecasts to the total of new World Bank lending. This step also allows the user to force changes in lending to specific countries via an exogenous multiplier (*xwblfinm*).

$$XWBLNFIN_{r,t} = XWBLnFinAdj_{r,t} * \frac{XWBLOANSTOT_t}{\sum^R XWBLnFinAdj_{r,t}} * xwblfinm_r$$

The outstanding loan portfolio for each country can then be updated.

$$XWBLOANS_{r,t} = XWBLOANS_{r,t-1} + XWBLNFIN_{r,t} - XWBLNFOUT_{r,t} - XWBSUBF_{r,t}$$

The third decision of the World Bank as agent was specified to be the target sector of loans within recipient countries. In large part because of the other representations within IFs, five such targets are specified: education, health, unskilled worker households, skilled worker households, and other uses. At this point in model development a single parameter (*xwbsector*) has been added with default values that draw upon data gathered about past Bank practice to guide flows to uses within recipient countries. A final adjustment term is put unto that equation based upon algorithmic analysis in the first year; for instance, if it is found that supposed World Bank lending to the education sector exceeds government spending in that sector, the adjustment factor reduces it.

$$XWBFLOWS_{r,ft,t} = XWBLNFIN_{r,t} * \frac{xwbsector_{t,ft}}{100} * XWBFlowsAdj_{r,ft,t=1}$$

The fourth and final set of decisions concern the terms of the lending. Again, parameters exist to change interest rates (*xwblintr*) and repayment periods (*xwblnrepr*). Currently, returns of funds to the World Bank and the IMF are assumed in the base case to be at the rate of 5% of principal each year, with 3% real interest payments, about the rate of global economic growth.

The IMF representation is fundamentally parallel to that for the World Bank. It determines in the same manner how much credit the IMF has to offer. And it determines the initial target of credits based on the scaling up of initial credits, the process that computed on a base calculation for the World Bank. There is no representation of IMF agency based on analysis around the drivers of its credit offering over time (such as liquidity problems of recipients); instead credit levels are now changed over time with GDP of recipients and an exogenous multiplier (*ximfcrfinm*). Agency could be developed analogously to the representation in the World Bank, although the real-world IMF process is more politically charged than are Bank lending decisions.

$$XIMFCREDITTOT_t = XIMFCREDITTOT_{t-1} + XIMFCRRepayTot_t - XIMFCRNewTot_t + XIMFCrNewSubTot_t * ximfcreqr$$

The three terms that adjust total IMF credits across years begin with repayments, the sum of flows out of recipient countries (XIMFCRFOUT) to pay either interest (at *simfcrintr*) or principal (at *ximfcrrepr*).

$$XIMFCRRepayTot_r = \sum^R XIMFCRFOUT_r$$

where

$$XIMFCRFOUT_r = XIMFCREDIT_{r,t-1} * \frac{simfcrintr}{100} + \frac{XIMFCREDIT_{r,t-1}}{ximfcrrepr}$$

The second term specifies the flow of subscriptions by donor countries to the IMF (XIMFSUBF) and is determined by the overall growth in IMF credits (*ximfcreditr*).

$$XIMFCRNewSubTot_t = \sum^R XIMFSUBF_r$$

where

$$XIMFSUBF_r = XIMFSUBF_{r,t-1} * \left(1 + \frac{ximfcreditr}{100} \right)$$

The third term carries the new credits provided to recipients (XIMFCRFIN). This is an initial computation to obtain the overall level of global IMF credits. Country-specific values will be recomputed later, but normalized to this overall level.

$$XIMFCRNewTot_t = \sum^R XIMFCRFIN_t$$

where

$$XIMFCRFIN_r = XIMFFLOWSINTOT_{t=1} * \frac{XIMFCREDITTOT_{t-1}}{XIMFCREDITTOT^1} * \frac{XIMFCRFIN_{r,t=1}}{XIMFFLOWSINTOT_{t=1}}$$

The next step is the reallocation of flows into countries. Instead of the more sophisticated agency representation of the World Bank, credit inflows are adjusted from initial value by growth in the potential GDP and by an exogenous multiplier (*ximcrfinm*). After an initial computation, the values are then normalized across countries so that the total global flows computed above still hold.

$$XIMFCRFInComp_r = \frac{XIMFCRFIN_{r,t=1}}{GDPTOT_{r,t=1}} * GDPTOT_r * ximcrfinm_r$$

$$XIMFCRFIN_r = XIMFCRFInComp_r * \frac{XIMFCrNewTot}{XIMFCRCompTot}$$

where

$$XIMFCRCompTot = \sum^R XIMFCRFInComp_r$$

Given the final calculation of IMF credit inflows, it is possible to compute the total IMF credit position (XIMFCREDIT) of each country (which will be negative for subscribers).

$$XIMFCREDIT_{r,t} = XIMFCREDIT_{r,t-1} + XIMFCRFIN_{r,t} - XIMFCRFOUT_{r,t} - XIMFSUBF_{r,t}$$

There were many data-related issues and decisions made to initialize the World Bank and IMF representations. The WDI database provides information on annual loans from the World Bank in two categories, those of the International Bank for Reconstruction and Development (IBRD) and those from the International Development Bank (IDA). The latter loans tend to have more concessional terms than do the former. Similarly, the database provides information on annual credits from the IMF, dividing them into concessional and non-concessional.

For the purposes of the SAM at this stage, the pre-processor sums the two types of loans and two types of credits into a total annual lending by the Bank and total credit flow from the IMF. Because these values tend to fluctuate substantially from year to year, and because the final years of the 1990s were a rather unusual period in global financial markets, the values set for initialization of forecasts are calculated from 5-year averages of loans and credits received as a percent of GDP in recent years, with the average rate then applied to the GDP of developing countries in the model's base year.

Unfortunately, the data only indicate net flows. For purposes of the dynamics of the model, it is important to distinguish repayment of loans or credits and interest on loans from new inflows. Net flows from the World Bank were decomposed into estimates of inflows (XWBLENFIN) and outflows (XWBLENFOUT) by estimating the value of repayments (outflows) based on the stock of loans from the Bank; initialization of stocks will be discussed below. The estimate of outflows assumed a 20-year life on loans and therefore 5% annual repayment, as well as a 3% interest rate via parameters that have now been made flexible for intervention by users. New inflows are then calculated as net flows minus the outflows. Exactly the same process was followed for IMF credits.

The pre-processor turns next to lending of the World Bank and credits from the IMF, collectively the International Financial Institutions (IFIs). It again uses WDI data for initial conditions of total loans held by borrowing countries (XWBLOANS). Nulls are filled with zeros. The same is done for IMF credits (XIMFCREDIT). In both cases, it is assumed that the bulk of funds loaned by the two IFIs come from subscriptions that ultimately constitute assets of developed countries. The pre-processor assigns assets by size of GDP.

3.3.13 International Accounting and Financial Equilibration

The documentation of the domestic side of the SAM representation explains equilibrating dynamics of two kinds: those around goods and service markets and fluctuations in inventories or stocks (both by sector via relative prices and overall via interest rates) and those around financial assets and liabilities of agent-classes (via adjustment multipliers sensitive to the level of debt), especially government debt. On the international side of the SAM structure, the key equilibrating dynamic centers on the total level of international debt of a country. Changes in foreign reserves and short-term governmental borrowing play important short-term buffering roles. Exchange rates are the key longer-term equilibrating variable.

Conceptual foundations for the dynamics build largely on two concepts: current account balance (CURACT) and capital account balance (CAPACT). Although in terms of definition they are equal with opposite signs, it is more complicated than that. Current account equals the trade balance (TRADEBAL) of relative-price adjusted exports (XRPA) minus imports (MRPA), plus the net outflow of foreign aid (AID), plus foreign workers' remittances to home countries, plus pension payments to retirees living abroad, plus the rents, interest, profits, and dividends paid on past capital inflows. Capital account equals the net inflow of underlying capital flows (the change in stocks), both long-term and short-term. The long-term includes foreign direct investment; the short term includes hot money in bank deposits and T-bills.

Although the current and capital accounts must balance, the balance often relies in the very short term on residual changes in stocks, namely reserve holdings and reactive government borrowing. In the longer-term exchange rate and interest rate fluctuations help maintain the balance by affecting the dynamics of the more slowly adjusting trade and longer-term financial flows. IFs brings together most of these elements in its representations of the international financial dynamics around the SAM.

3.3.13.1 Current Account

IFs computes the current account balance from the elements indicated above, except that pension liabilities abroad are not represented. Although actual foreign direct investment and portfolio flows belong in the capital account (see below), the net earnings on the assets abroad and in-country belong in the current account. In the formulation below, not all government or firm debt is represented as a basis for generating international flows, only international government debt (XGOVTDEBT, not GOVDEBT, which is also different in that it in IFs it is a percent of GDP) and only international firm debt (XFIRMDEBT).

$$Curact_r = XRPA_r - MRPA_r + AID_r + XWORKREMIT_r - Interest_r$$

where

$$Interest_r = \mathbf{lintr} * (IFIDebt_r + FDIDebt_r + PortDebt_r + FirmDebt_r + GovDebt_r)$$

$$IFIDebt_r = XWBLOANS_r + IMFCREDIT_r$$

$$FDIDebt_r = XFDISTOCK_r - XFDISTOUT_r$$

$$PortDebt_r = XPORTFOLIO_r - XPORTSTOUT_r$$

$$FirmDebt_r = XFIRMDEBT_r$$

$$GovtDebt_r = XGOVTDEBT_r$$

Most of the above formulation is focused on determining the net interest, profit, dividend flows, which are specified as an exogenous real interest rate (*lintr*) times the net stocks of assets mostly discussed earlier.

3.3.13.2 Capital Account

IFs computes a limited (or partial) capital account balance from the elements of the capital account over which the government normally has least direct control. The limited capital account balance purposefully excludes the two equilibrating elements of changes in reserves and short-term government compensatory borrowing.

$$CapActLim_r = FDIFlow_r + PortFlow_r + IFIFlow_r$$

where

$$FDIFlow_r = XFDIFIN_r - XFDIFLOUT_r$$

$$PortFlow_r = XPORTFIN_r - XPORTFOUT_r$$

$$IFIFlow_r = XWBLNFIN_r - XWBLNFOUT_r + XIMFCRFIN_r - XIMFCRFOUT_r$$

3.3.13.3 External Debt

The current and capital accounts come together in the determination of external balancing debt. The model makes no assumption that it can anticipate a government's relative choice between changes in reserves and changes in debt taken on for balancing purposes (XGOVTDEBTB) to satisfy the discrepancy between the current and capital accounts. . In addition to updating current year values of that debt with current and capital account information (in IFs the sign on capital account is reversed from that on the current account, so both are subtracted), the loan portion of foreign aid (LoanPortionAid), determined by the parameter *aidlp*, is also added to the debt.

$$\begin{aligned} XGOVTDEBTB_{r,t+1} \\ = XGOVTDEBTB_{r,t} - CURACT_r - CapActLim_r + LoanPortionAid_r \end{aligned}$$

where

$$LoanPortionAid_r = AID_r * aidlp_r$$

Total external governmental debt is set as the basic debt plus (normally longer-term) IFI credits and loans.

$$XGOVTDEBT_{r,t+1} = XGOVTDEBTB_{r,t+1} + XIMFCREDIT_{r,t} + XWBLOANS_{r,t}$$

Total external debt is the sum of governmental indebtedness and firm debt.

$$XDEBTRPA_{r,t+1} = XGOVTDEBTB_{r,t} + XFIRMDEBT_{r,t}$$

Although IFs does not separate out reserve position changes from this debt, it has begun to develop an initial representation of reserves (XRESERVES) and at some point that separation may occur. Currently the equations (not shown) simply move XRESERVES from the initial portion of GDP toward 15 percent of GDP. That allows a quick calculation of liquidity as the current account ratio to the reserves.

$$LIQUIDITY_r = \frac{CURACT_r}{XRESERVES_r} * 100$$

3.3.13.4 External Debt for exchange rates and external accounts balancing

Exchange rates are a function of external debt, using a PID controller mechanism to maintain relative balance with a target over time of "TypicalDebt" at a given GDP per capita level. The target can be modified with an exogenous parameter (*repayrm*).

$$XDebtTargetPerGDP_r = ConvergeOverTime(XDebtIRat_r, UltimateTarget_r, 100)$$

where

$$XDebtIRat_r = \frac{XDEBTRPA_{r,t=1}}{GDPPOT_{r,t=1}}$$

$$UltimateTarget_r = Amin(XDebtIRat_r, TypicalDebt_r) * repayrm_r$$

$$TypicalDebt_r = F(GDPPCP_r)$$

The exchange rates change with a multiplier computed from the debt levels, where the multiplier and therefore the change use the PID adjuster function of IFs so as to control movement toward equilibrium with both the absolute debt level and the change in it. The exchange rate term is also imbedded in a moving average formulation, pulling it very slowly back toward its initial index value of 1.0.

$$EXRATE_r = EXRATE_{r,t-1} * Mul * 0.98 + .02 * 1.0$$

where

$$Mul = ADJSTR(GDPPOT_r, Diff1, Diff2, Par1, Par2)$$

$$Diff1 = XDEBTRPA_r / GDPPOT_r - XDebtTargetPerGDP_r$$

$$Diff2 = XDEBTRPA_r / GDPPOT_r - XDEBTRPA_{r,t-1} / GDPPOT_r$$

The parameters used in the exchange rate adjustment above are second-order parameters computed from exogenous specifications (*elerca1*, *elerca2*). For entrepot economies with a high ratio of trade to GDP, the parameters are scaled back so as to avoid over-reaction.

$$Par1 = elerca1 * mul * 0.5$$

$$Par2 = \mathit{elerca2} * \mathit{mul} * 0.35$$

where if

$$X_{r,t-1} + M_{r,t-1} > GDPPOT_r$$

then

$$\mathit{Mul} = \frac{GDPPOT_r}{X_{r,t-1} + M_{r,t-1}}$$

else

$$\mathit{Mul} = 0.2$$

3.3.13.5 Foundations in the preprocessor for extending treatment of debt

With debt, initialization of concepts with data was a challenge. The first step in the initialization of all financial stocks within the pre-processor is the processing of data on total external debt of countries (XDEBTRPA). The raw data came from the WDI. Null values for developing countries are filled with zeros. Unfortunately, external debt data from the WDI are for developing countries only and there are no corresponding asset data for the lending countries, presumably mostly developed countries. As in other such instances, the pre-processor now assigns the total of external debt of developing countries to developed countries, defined for this purpose as countries with no external debt indicated and GDP per capita of \$10,000 or more. The assignment is proportional to GDP of the more developed countries. Both the presumption that more developed countries are primarily asset holders and the proportional assignment are clearly inaccurate. The rationale for both is that the resultant stock assumptions, on average, are almost certainly better than assumptions that asset stocks of such countries are uniformly zero. Further, the dynamics of the model will be more greatly determined by change in stocks from initial levels than by the absolute values of them. Nonetheless, data on these assets would be highly desirable.

In the preprocessor some steps have been taken to further refine the specification of government debt, for instance the separation of externally-held debt from domestically-held debt and the differentiation of publicly guaranteed and not-guaranteed debt. Although these distinctions have not been brought into the forecasting side of IFs yet, and the need for them is not great, it merits comment on the preprocessor foundations.

With respect to firms, the preprocessor reads data from WDI on publicly guaranteed and not-guaranteed debt, both of which are assumed to be held by firms and which are summed to obtain total external firm debt (XFIRMDEBT). Holes are filled with zeros because levels of debt are often very small and non-reporting often will imply debt levels near zero. Because the data only cover stocks of debt held by developing countries as liabilities, the stocks as assets were once again allocated to developed countries, specifically to firms within them, according to GDP size.

With respect to government, not all of government debt is external. In fact, especially for developed countries, households and firms often hold much of it. In order to divide government debt between foreign and domestic holders of it, the preprocessor first computes government external debt. This is done by assuming that households do not have any real external debt, so that all external debt not assignable to firms is governmental external debt.

$$XGOVDEBT_r = XDEBT_r - XFIRMDEBT_r$$

The assumption was again made that all of this debt should be assigned as assets to governments in developed countries. Rather than assigning it on the basis of GDP, however, it was assigned on the basis of official development assistance (ODA). The OECD and WDI both provide data on ODA. Because significant portions of ODA come as loans rather than grants, the assignment of the government asset offset of developing government debt based on the size of ODA makes sense.

Again, only in the preprocessor and not in the forecasting portion of the model, governmental domestic debt is calculated as total government debt minus the external stock.

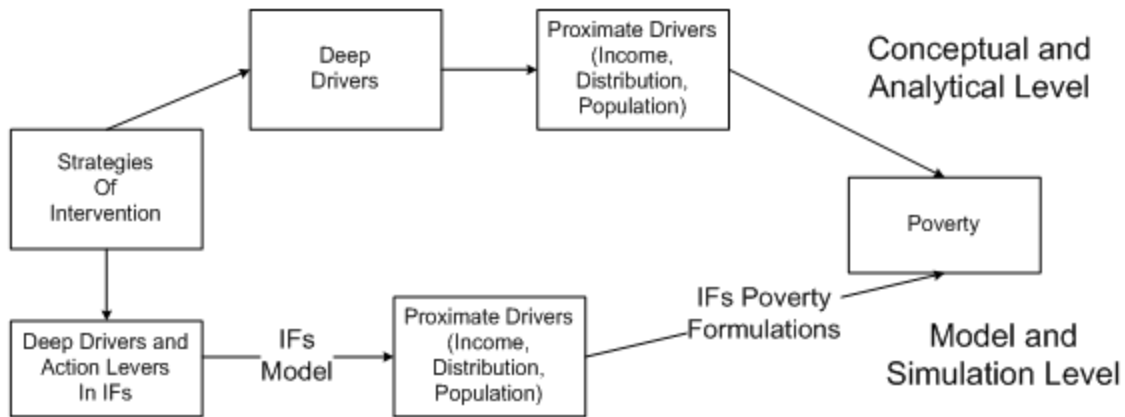
$$GOVTDEBTD_r = TotGovtDebt_r - XGOVDEBT_r$$

3.4 Poverty

The IFs project has given special attention to the forecasting of poverty, most notably in the first volume of the project's Patterns of Potential Human Progress series, titled *Reducing Global Poverty* (Hughes et al. 2009). We therefore want to document the approach taken within IFs to that forecasting. We begin with the conceptual foundations and turn to the specific forecasting formulations.

3.4.1 *The Proximate Drivers of Poverty*

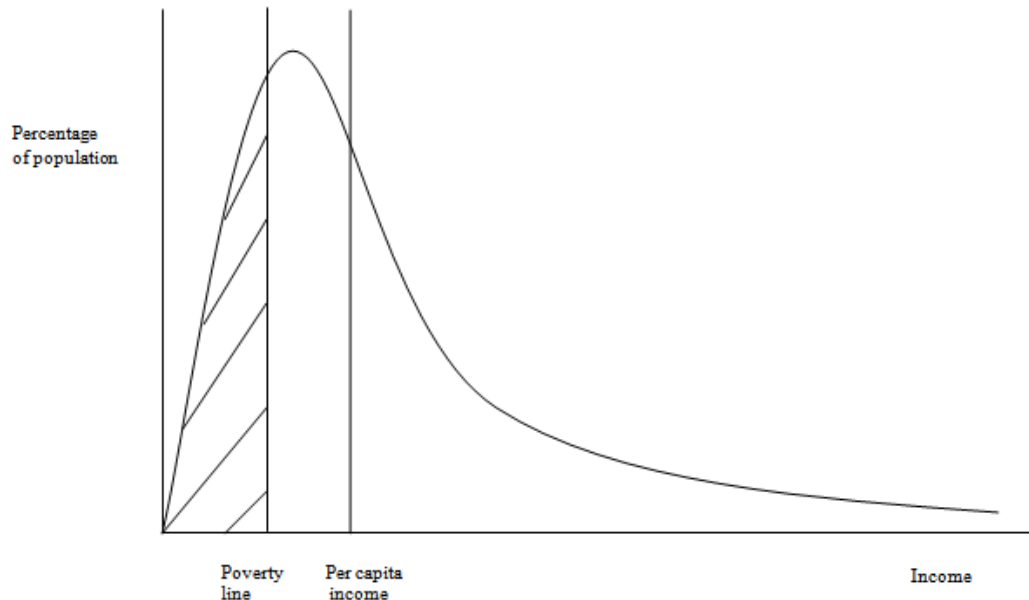
Knowledge of two drivers allows quite accurate calculation of the portion of a population living at or below any specified poverty level: the average income of the population and the distribution of that income. Once the rate of poverty is known, it requires only knowledge of a third driver, population size, to know also the number of people living below that poverty level. In dynamic societies there are, of course, a huge number of deeper or more distal drivers that affect these proximate drivers (literally anything that affects economic growth, the nature of income distribution, and the population size). The figure below shows this conceptual understanding of poverty, and indicates that the IFs system formalizes that understanding by calculating poverty from the three proximate drivers and also calculates those proximate drivers from the much larger forecasting system (the economic, population, and other models of IFs). Moreover, the IFs system makes available large numbers of intervention points for developing alternative scenarios around poverty futures.



Other parts of the documentation of IFs describe the determination of population and economic growth and of the calculation of the Gini coefficient (GINIDOM) as the key representation within IFs of income distribution. That documentation describes how the Gini coefficient is based on the Lorenz curve and how the Lorenz curve in IFs is tied to forecasting of the income shares of unskilled and skill households and the shares of population in those two household types.

The Lorenz curve of IFs with only two income categories is, however, an unsatisfactory form for representing the income distribution of the societies we forecast. It is both “non-parametric” in the sense that no simple analytical form can describe it for the purposes of our forecasting and it is crude in the representation of just the two income types. What we want for poverty forecasting is an analytical form and one that can capture more smoothly the increasing income shares that richer portions of the population will have. (Ideally we would also want a representation from which we can conveniently compute specific deciles or quintiles.)

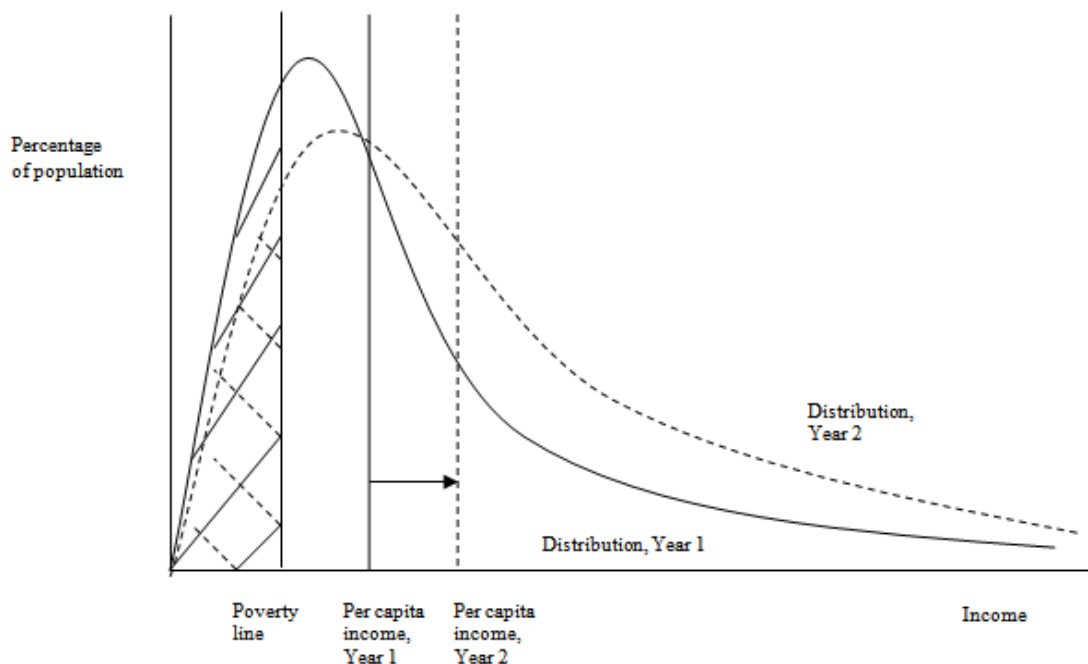
The most widely used parametric representation in poverty analysis is the lognormal density. A **density curve** captures the percentage of the population that earns or consumes a given amount (unlike a distribution that captures the cumulative percentage of population that earns or consumes up to a given amount). Although income and consumption are not exactly distributed in a lognormal form for every country, it is a very good approximation to observed empirical distributions for most (it will obviously not be such a good form for a country like post-Apartheid South Africa that has something closer to a binormal distribution). As Bourguignon (2003: 7) notes, a **lognormal distribution** is “a standard approximation of empirical distributions in the applied literature.” A variable is lognormally distributed if the natural logarithm of that variable is normally distributed, as in the well-known “bell curve.” See the figure below a lognormal density curve.



One advantage of using a lognormal density to capture the distribution of income in a society is that it can be fully specified with only two parameters, average income and the standard deviation of it. More useful for our purposes, the Gini coefficient can be used in lieu of the standard deviation. See Appendix 2 of Hughes et al. (2009) for the explanation of that relationship and for much more information on the use of the lognormal distribution in IFs, including the calculations of poverty rates and poverty gaps, and the reconciliation of analysis using national accounts and survey data.

The figure above provides an illustration of how to obtain the poverty headcount from a lognormal density curve. For a specified poverty line—for example, the one corresponding to dollar a day—the area to the left of the line gives the poverty headcount ratio. The first vertical line in the figure shows the poverty line, and hatched lines show the area corresponding to the headcount ratio. The poverty headcount *number* is the headcount *ratio* times population.

The figure below similarly illustrates how changes in income distribution and the average level of per capita income can affect the poverty rate and therefore number. In this figure, the second vertical line from the left shows the income level for a given point in time, say Year 1, while the solid curve shows the income distribution. The third and dashed vertical line shows the income for a subsequent time, say Year 2, while the dashed curve shows an associated income distribution. The area of hatched, dashed lines to the left of the poverty line in the new distribution shows the new poverty headcount ratio. This area is smaller than the area under the Year 1 distribution, and the poverty headcount ratio has decreased. What happens to the poverty headcount number depends on how the population has changed between the two years. If the population increases significantly, the headcount number can increase even if the headcount ratio decreases.



What is the evidence on poverty changes arising from the interaction of growth and distributional effects? There is evidence that growth tends to be “distribution neutral” on average; Ravallion and Chen (1997), Ravallion (2001), Dollar and Kraay (2002), find a close to zero correlation between changes in inequality and economic growth. This is consistent with the evidence that the growth effect dominates and growth tends to reduce absolute poverty (World Bank 1990, 2000; Ravallion 1995; Ravallion and Chen 1997; Fields 2001, and Kraay 2004). World Bank (2001) and Ravallion (2004) suggest that the “elasticity” of the dollar-a-day poverty rate to growth is -2 ; an increase in the growth rate by 1% is associated with a decrease of 2% in the headcount index of poverty.

While there is general consensus that growth is good for poverty alleviation, a few voices of caution can be heard. The actual reduction in poverty is arguably lower than might be expected given recorded rates of economic growth. This has been termed “the paradox of persistent global poverty” (Cline 2004: 28). Poverty in the 1990s declined by less than would have been predicted with a poverty-growth elasticity of around -2 . Khan and Weiss (2006) warn that the elasticity of poverty to growth can vary widely—only -0.7 for the Philippines compared to -2.0 for Thailand—depending on the initial inequality and changes in inequality over time. Ravallion (2004) lists a few reasons to be cautious about the finding of distribution neutrality of growth: measurement error in changes in inequality, possible churning under the surface with winners and losers at all income levels, and possible increases in absolute income disparities. Moreover, a few countries and regions could experience poverty increases from distributional changes, even if on average there is neutrality.

In addition to uncertainties introduced by income inequality effects, the elasticity approach to anticipating poverty decline with income suffers from another problem. A given rate of economic growth will have a bigger impact on poverty headcount when the

poor are clustered closely around the poverty line than when their incomes fall markedly below the line. In some countries this phenomenon might explain the weaker than expected response of poverty levels to growth in the face of only modest changes in overall inequality.

The lognormal approach for forecasting poverty of IFs eliminates the need to use the elasticity approach. In fact, lognormal specifications could be used to calculate variable poverty elasticities across countries and time.

3.4.2 Lognormal Analysis of Change in Poverty: IFs Formulations

Given its advantages, the IFs approach to forecasting poverty uses the lognormal formulation, driven by average income (actually household consumption) and the Gini coefficient. One advantage of the approach is that the two drivers (average consumption and the log-normal function/distribution linked to Gini) allow the computation of poverty rates at any poverty level specified and a specialized poverty display within the IFs system allows the user to do just that. Yet to be consistent with the Millennium Development Goals and the common discourse we focus on levels such as \$1.25 and \$2.00 per day, specified in terms of purchasing power. The computed variables in IFs use those levels (INCOMELT1LN and INCOMELT2LN).

Although it is common to speak of poverty in terms of income, much of the survey data used to identify historical and contemporary levels has focused on consumption and the IFs calculations are driven by household consumption per capita (C divided by POP). Another and related important decision was the use of changes of household consumption to drive poverty over time—although survey data from the World Bank's PovcalNet database are used to initialize poverty rates and numbers, C and domestic income distribution (GINIDOM) change values over time. It is well-known that calculations from national accounts data tend to provide lower poverty rates than those found in surveys. IFs computes shift factors in the first year that identify the ratio of values from surveys to those from national accounts (NSNARAT) and maintains those shifts over time (although an exogenous parameter, *nsnaratm* can adjust those over time). Another exogenous parameter can change the poverty level from \$1 (*incomepovlevm*).

The process in IFs begins with the calculation of the lognormal standard deviation (LNStDev) from Gini and with the calculation of the mean consumption (MeanConsumption) from consumption per capita and a smoothed conversion factor from market exchange rates to purchasing power parity (SmPPPConv). These are the key drivers of the lognormal formulation. See Appendix 2 of Hughes et al. (2009) for a more formal and extended explanation of the entire poverty calculation system using the lognormal approach.

$$LNStDev_r = Sqr(r) * CalcNormalInverse\left(\frac{GINIDOM_r + 1}{2}\right)$$

$$MeanConsumption_r = \frac{C_r}{POP_r} * SmPPPConv_r$$

where

$$SmPPPConv_r = SmPPPConv_r * 0.8 + PPPCONV_r * 0.2$$

Given these driving inputs, IFs uses a two-step process for calculating the share of population below any given consumption level. The first step produces a log normal mean (LNMean) based on the mean consumption adjusted by the national accounts scaling factor from the first year (NSNARatFac) and the lognormal standard deviation.

$$LNMean_r = \text{Log}(\text{MeanConsumption}_r * NSNARatFac_{r,t=1} * nsnaratm_r * 1000) - .05 * LNStDev_r^2$$

where

$$NSNARatFac_{r,t=1} = \text{Back Calculation from } LNMean_{r,t=1} \text{ and } \text{MeanConsumption}_{r,t=1}$$

Finally, a specialized function computes the rate of poverty to be multiplied by the population to get the poverty numbers.

$$INCOMELT1LN_r = \text{CalcLogNormalCDF}(365 * incpovlevm, LNMean_r, LNStDev_r) * POP_r$$

The calculation of poverty below \$2 is the same, except that the value of 365 in the above equation becomes $365 * 2 = 730$.

In 2014 the International Comparison Project (ICP) rebased its estimates for all countries of GDP at purchasing power from 2005 values and dollars (at PPP) to 2011 values and re-estimated PPP values. In order to pick up new estimates of income poverty, IFs added two new variables, one for poverty at \$1.25 in 2011 PPP dollars (INCOMLET1LN2011) and another similarly for poverty at \$0.70 (INCOMELT07LN2011). It was necessary to re-estimate the appropriate value of household consumption (C) for these estimates. In the sequence of calculations above this was possible with the reformulation below (and related changes for the \$0.70 calculation):

$$\text{MeanConsumption}11_r = \frac{C_r}{POP_r} * \text{RatioGDP2011to2005ICP}_{r,t=1} * \text{SmPPPConv}_r$$

where

$$\text{RatioGDP2011to2005ICP}_{r,t=1} = \text{GDPPCP2011}_{r,t=1} * \text{GDPPCP2005}_{r,t=1}$$

The calculation of the poverty gap at \$1.25 (POVGAP) relies on the same basic lognormal structures and some of the same input variables. The poverty gap is defined as the mean shortfall from the poverty line, expressed as a percentage of the poverty line; the non-poor are considered to have zero shortfall. The measure helps identify the depth of poverty and is therefore a useful supplement to its incidence.

$$POVGAP_r = \text{PortionPoor}_r - \frac{\text{MeanBackCalculation}_r}{365 * incpovlevm} * \text{IncFrac}_r * 100$$

where

$$\text{PortionPoor}_r = \frac{INCOMELT1LN_r}{POP_r}$$

$$MeanBackCalculation_r = Exp(LNMean_r + 0.5 * LNStDev_r^2)$$

3.4.3 Cross-sectional Analysis of Change in Poverty

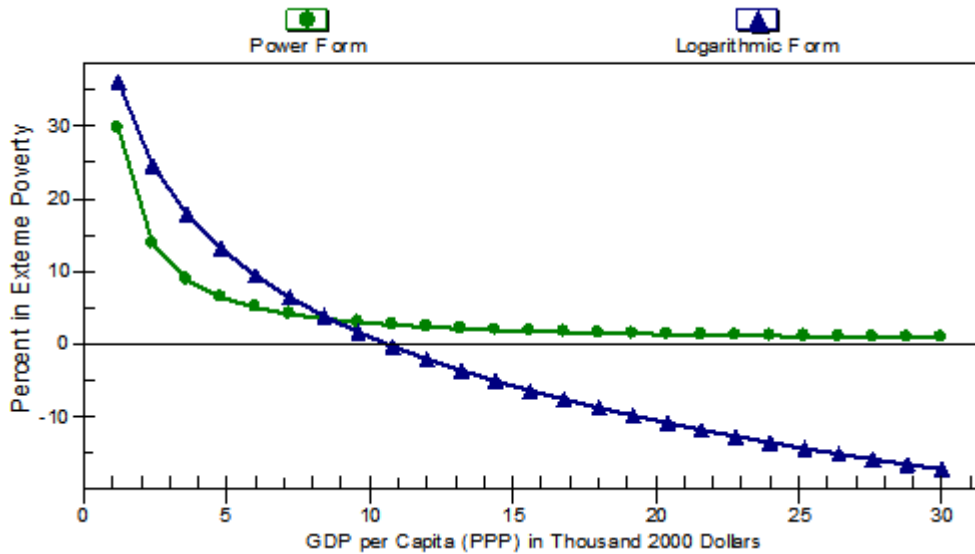
Although the dominant approach to forecasting poverty in IFs used the lognormal formulation, a simpler approach is also possible using a cross-sectional estimation of the relationship between poverty and GDP per capita at PPP. This simpler approach was developed for IFs before the lognormal one and has been retained as a useful second estimate of future poverty. In addition, in the pre-processor the cross-sectional formulation is used to fill holes for countries without surveys in estimating initial poverty rates (even when forecasts then go on to use the logarithmic approach). Not surprisingly, the cross-sectionally estimated function uses the same proximate driving variables as the lognormal function, namely GDP per capita at PPP (GDPPCP) and domestic Gini (GINIDOM). A scaling factor computed the first year (IncomeLT1RI) assures that the value computed in the first year matches data; the factor is maintained in the calculation over time. The formulation for poverty at \$2.00 per day (INCOMELT2CS) uses the same formulation.

$$INCOMELT1CS_r = AnalFunc(GDPPCP_r, GINIDOM_r) * IncomeLT1RI_{r,t=1}$$

where

$$IncomeLT1RI_{r,t=1} = \frac{INCOMELT1CS_{r,t=1}}{AnalFunc(GDPPCP_{r,t=1}, GINIDOM_{r,t=1})}$$

As with other functions used in IFs, we experimented with many variations before determining which to use in the model. The figure below shows the relationship between GDP per capita at PPP and poverty using both an exponential or power curve and a logarithmic form; both produce approximately the same R-squared. The logarithmic form is normally used in IFs, because it appears visually to capture better both the higher levels of poverty at low levels of GDP per capita and to capture the near elimination of extreme poverty by about \$10,000 per capita.



3.5 Informal Economy

We conceive of the informal economy as comprising two parts, informal production *and* informal labor. For informal production, we use the official adjustments to national accounts countries have made using the OECD’s non-observed economy (NOE) framework (OECD 2002). While this measure does sometimes include illegal production, household production for own-final use, and other activities not reported due to statistical errors or deficiencies, their contribution to the overall NOE is small compared to underground and informal production,. Although conceptually this aligns fairly closely with the Shadow Economy concept used by Schneider and Enste (2000) and others, the values produced by Schneider and Enste are produced by model estimation rather than survey methods and can be considerably higher than those of the OECD. We have added a variable to forecast those shadow economy values also, but in our Base Case do not use it in our forward linkage analysis.

For informal labor, we follow the lead of the ILO and WIEGO to include not just informal labor employed within informal enterprises or self-employed, but also informal labor in the formal sector and (paid labor) in households (ILO 2013; Chen 2012). The NOE framework assumes that informal labor in the formal sector is fully accounted for by standard NSA reporting practices, but we feel this assumption obscures an important aspect of the overall informal economy.

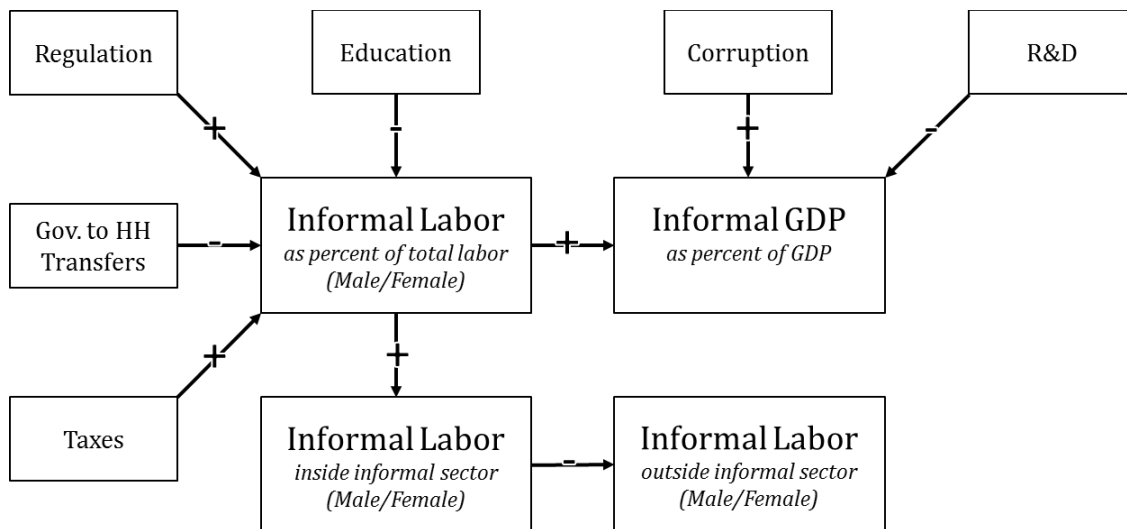
Based on this conception, we developed four statistical models for use in our forecasts of the informal economy: (1) a model of the size of the informal labor force; (2) a model of the distribution of informal labor residing either inside, or outside of the informal sector; (3) a model of the informal economy expressed as a percent of official GDP; and (4) a model of the shadow economy expressed as a percent of official GDP. We use the output of the first model as a driving variable for the subsequent three models.

3.5.1 Drivers of Informality

Across the three models identified above are seven explanatory variables of informality, each represented in the IFs model, and designed to both capture the long-term relationship between human development and informality as well as the drivers that relate to the motivations and choices of firms and individuals and to the levers or policies of governments. These drivers can be clustered into the following categories:

1. *Developmental drivers* are the root cause of involuntary informality, which is primarily low human capital. In developing countries with higher levels of survivalist informality, developmental drivers may dominate. Even in more developed economies, where informality tends to exhibit different characteristics, a deeper developmental driver may retain importance through its impact on the secondary drivers of informality.
2. *Secondary drivers* are the push and pull factors that affect whether an actor will chose to enter, remain in, or opt-out of the formal economy.

In the IFs model, *developmental drivers* are primarily represented by years of educational attainment by adults (15 years and older) which serves as a proxy for overall levels of human capital, while the push factors of the *secondary drivers* are represented in the model by the nature of the business regulatory environment, the size of the tax burden, and the level of corruption, and the pull factors by government to household transfers and spending on research and development.



Drivers of informality in IFs

Source: authors' conception

3.5.2 Forecasting Informal Labor Share

The IFs model provides an estimate of the overall size of the informal economy by calculating the shares of total labor and official GDP that are informal. The model begins by calculating the share of informal labor. The forecast of informal labor share as a

percentage of total labor force (LABINFORMSHR) is computed in two primary steps. First, a basic value of that percentage (LabInformal) is calculated as a function of four drivers: (1) Years of adult education (15+) (EDYRSAG15); (2) Government Business Regulation Index (GOVBUSREGIND); (3) Firm Tax Rates (FIRMTAXR); and (4) Government Transfers to Households for Welfare and Pensions as a percent of GDP (GOVHHTRN/GDP*100). The four driver terms are all from the previous model year (t-1), because in the recursive structure of IFs they are computed later in the sequence of variables.

$$\begin{aligned}
 LabInformal_{r,t} = & \mathbf{labinformcoeffintercept} \\
 & + \mathbf{labinformcoeffed} * EDYRSAG15_{r,t-1,s=3} \\
 & + \mathbf{labinformcoeffbus} * GOVBUSREGIND_{r,t-1} \\
 & + \mathbf{labinformcoefftax} * \text{Log}(FIRMTAXR_{r,t-1}) \\
 & + \mathbf{labinformcoeffhhtrn} * \left(100 * \frac{GOVHHTRN_{r,t-1}}{GDP_{r,t-1}} \right)
 \end{aligned}$$

In the above equation, the basic values of the parameters are: $\mathbf{labinformcoeffintercept} = -12.613$, $\mathbf{labinformcoeffed} = -2.510$, $\mathbf{labinformcoeffbus} = 12.604$, $\mathbf{labinformcoeffhhtrn} = -1.810$, $\mathbf{labinformcoefftax} = 10.6$. And in the above equation, education years are for the total population (sex=3) rather than males or females.

The second step modifies that initial calculation by converging a base year shift term (LabInformShrShift) to zero over years specified by $\mathbf{labinfconvyr}$ and by applying an exogenous multiplier ($\mathbf{labinformshrm}$).

$$\begin{aligned}
 LABINFORMSHR_{r,t} \\
 = & (LabInformal_{r,t} \\
 & + \mathbf{ConvergeOverTime}(LabInformShrShift_{r,t=1}, 0, \mathbf{labinfconvyr}_{r,t})) \\
 & * \mathbf{labinformshrm}_{r,t}
 \end{aligned}$$

where

$$LabInformShrShift_{r,t=1} = LABINFORMSHR_{r,t=1} - LabInformal_{r,t=1}$$

and

$$\begin{aligned}
 \mathbf{labinfconvyr}_{r,t} & = 60 \text{ if } LabInformShrShift_{r,t=1} > 0 \\
 \mathbf{labinfconvyr}_{r,t} & = 120 \text{ if } LabInformShrShift_{r,t=1} \leq 0
 \end{aligned}$$

Note that, in the above formulation, an adjustment to the calculation of LABINFORMSHR is required to smooth the transition from historical data to forecast values, because the initial (t=1) value of that variable from data will not be consistent with the calculation of it in the function that generates LabInformal. The adjustment, or shift factor (LabInformShrShift), is the difference in the first year between the value of LABINFORMSHR and the calculated value from the driver function. The shift factor constrains the model output in the initial year to better match historical data and then gradually lets the forecasting formulation take over. It would be possible to maintain that additive shift factor as a constant over time, but because both informal labor shares generally decline with economic growth, we converge that factor over time toward 0 or lack of adjustment. If the shift factor is greater than 0, we converge it in 60 years;

otherwise in 120 years. The reason for the slower upward convergence is that, in general, the informal labor share decreases with development and advance in GDP per capita; upward movement would be very unusual.

Convergence parameters are added to the model when data values for the base year (2010) of variables differ substantially from those that forecasting functions also provide for that base year. In this situation we compute shift factors that record the discrepancy in the base year. Although those shift factors could be maintained as constant adjustments to forecasting formulations over time, our experience is that the reasons for the discrepancies fall into 2 categories: (1) errors in data that we have no ability to eliminate and (2) influences of variables not in our forecasting formulation, such as country-specific historical path dependencies related to historical events or cultural patterns. Over time data errors tend to be corrected and country-specific influences tend to erode, suggesting that we should generally converge the shift factors toward 0 over time. Because both bases for the shift factor involve information that we cannot have, the convergence time is a parameter that must be set subjectively. Our basis for doing so is analysis of model behavior with longer or shorter shift factors, with an eye especially to avoiding transients in behavior by having convergence times that are too short. Thus most convergence times are set at 50 years or more, with such attention to their implications for model behavior. In cases where we believe that shifts are most likely to occur in one direction (e.g. we expect more movement downward in shares of economies that are informal than upward), we will set the convergence times differentially as in the equation indicated.

3.5.3 Forecasting Informal Labor by Sector

It is very useful in our forecasting to know more about the makeup of informal labor. In particular, we want to know more about where that labor is employed and how it breaks down by sex. With respect to where it is employed, we use the categorization developed by the International Labor Organization and WEIGO, namely the portion of informal labor employed outside the informal sector (in the formal sector or in households) and the portion employed informally. We further divide each sub-section by sex.

The first step in computation is of the portion of informal labor employed inside the informal sector ($LABINFORMPCNTINF$), which varies inversely with the total share of labor that is informal ($LABINFORMSHR$). That is, as the informal share of total labor rises, a smaller portion of the total informal labor will be able to find positions in the formal sector (or as the informal share of total labor falls, the portion of the remaining informal labor employed in the formal sector will rise). This is computed for both sexes ($s=Total$) prior to a subsequent division by sex.

$$LABINFORMPCNTINF_{r,s=Total,t} = \mathbf{ininformcoeffintercept} + \mathbf{ininformcoefflabinf} * LABINFORMSHR_{r,t}$$

The values of the parameters currently are: $\mathbf{ininformcoeffintercept}=55.243$; $\mathbf{ininformcoefflabinf}=- 0.415$.

The percentage of the informal labor that is outside the informal sector (LABINFORMPCNTNONINF) is simply the residual.

$$LABINFORMPCNTNONINF_{r,s=Total,t} = 100 - LABINFORMPCNTINF_{r,s=Total,t}$$

The further division by sex of the informal labor by sector is done by using the initial patterns by sex (LABINFORMPCNTINFFemShare and LABINFORMPCNTNONINFFemShare). There are not enough data and conceptual foundation to establish a dynamic relationship for the share by sex.

$$\begin{aligned} LABINFORMPCNTINF_{r,s=Female,t} &= LABINFORMPCNTINF_{r,s=Total,t} * \\ & \quad LABINFORMPCNTINFFemShare_{r,t=1}) \\ LABINFORMPCNTINF_{r,s=Male,t} &= LABINFORMPCNTINF_{r,s=Total,t} \\ & \quad - LABINFORMPCNTINF_{r,s=Female,t} \end{aligned}$$

where

$$LABINFORMPCNTINFFemShare_{r,t=1} = \frac{LABINFORMPCNTINF_{r,s=Female,t=1}}{LABINFORMPCNTINF_{r,s=Total,t=1}}$$

Similarly,

$$\begin{aligned} LABINFORMPCNTNONINF_{r,s=Female,t} &= LABINFORMPCNTNONINF_{r,s=Total,t} * \\ & \quad LABINFORMPCNTNONINFFemShare_{r,t=1}) \\ LABINFORMPCNTNONINF_{r,s=Male,t} &= LABINFORMPCNTNONINF_{r,s=Total,t} \\ & \quad - LABINFORMPCNTNONINF_{r,s=Female,t} \end{aligned}$$

where

$$LABINFORMPCNTNONINFFemShare_{r,t=1} = \frac{LABINFORMPCNTNONINF_{r,s=Female,t=1} LABINFORMPCNTNONINFFemShare_{r,t=1}}{LABINFORMPCNTNONINF_{r,s=Total,t=1}}$$

3.5.4 **Forecasting Informal GDP Share**

The IFs model uses four drivers to forecast informal GDP share: (1) the informal labor share (LABINFORMSHR)—see section 3.5.1; (2) the portion of informal labor employed inside the informal economy (LABINFORMSEC, informal)—see section 3.5.2; (3) the level of corruption (GOVCORRUPT); and (4) the society's expenditure on research and development as a percentage of GDP (RANDEXP).

As in the computation of informal labor share, the forecast of informal GDP share as a percentage of official GDP (GDPINFORMSHR) involves two major steps. First, a basic value of that percentage (GDPInformal) is calculated as a function of the four drivers identified above. The driver terms are all from the previous model year (t-1), because in the recursive structure of IFs they are computed later in the sequence of variables.

$$\begin{aligned} GDPInformal_{r,t} = & \mathbf{gdpinformcoeffintercept} \\ & + \mathbf{gdpinformcoefflabinf} * LABINFORMSHR_{r,t-1} \\ & + \mathbf{gdpinformcoeffcorrupt} * GOVCORUPT_{r,t-1} \\ & + \mathbf{gdpinformcoeffRandD} * RANDDEXP_{r,t-1} \end{aligned}$$

In the above equation, the basic values of the parameters are: $\mathbf{gdpinformcoeffintercept} = 18.861$, $\mathbf{gdpinformcoefflabinf} = 0.193$, $\mathbf{gdpinformcoeffcorrupt} = -1.585$, and $\mathbf{gdpinformcoeffRandD} = -2.6$. The negative value on the corruption term is potentially misleading because in Transparency International's measure, higher values are more transparent and lower ones are more corrupt; the equation term increases GDPInformal as corruption increases.

The second step modifies that initial calculation by converging a base year shift term (GDPInformShrShift) to zero over a number of years specified by $\mathbf{gdpinfconvyr}$ and by applying an exogenous multiplier ($\mathbf{gdpinformshrm}$).

$$\begin{aligned} GDPINFORMSHR_{r,t} \\ = & (GDPInformal_{r,t} \\ & + \mathit{ConvergeOverTime}(GDPInformShrShift_{r,t=1}, 0, \mathbf{gdpinfconvyr}_{r,t})) \\ & * \mathbf{gdpinformshrm}_{r,t} \end{aligned}$$

where

$$GDPInformShrShift_{r,t=1} = GDPINFORMSHR_{r,t=1} - GDPInformal_{r,t=1}$$

and

$$\begin{aligned} \mathbf{gdpinfconvyr}_{r,t} & = 70 \text{ if } GDPInformShrShift_{r,t=1} > 0 \\ \mathbf{gdpinfconvyr}_{r,t} & = 200 \text{ if } GDPInformShrShift_{r,t=1} \leq 0 \end{aligned}$$

Note that in the above formulation, an adjustment is required to the calculation of GDPINFORMSHR, because the initial (t=1) value of that variable from data will not be consistent with the calculation of it in the function that generates GDPInformal. The adjustment or shift factor (GDPInformShrShift) is the difference in the first year between the value of GDPINFORMSHR and the calculated value from the driver function. It would be possible to maintain that additive shift factor as a constant over time, but because both informal GDP shares generally decline with economic growth, we converge that factor over time toward 0 or lack of adjustment. If the shift factor is greater than 0, we converge it in 70 years; otherwise in 200 years. The reason for the slower upward convergence is that, in general, the informal labor share decreases with development and advance in GDP per capita; upward movement would be very unusual.

3.5.4.1 Shadow economy

Shadow economy as a share of total is calculated using a formulation similar to that used for the informal share of the economy. Estimations from Schneider and Elgin were used to initialize the model in the base year (countries with missing data are initialized using a regression of the most recent year data against per capita income of countries). The regression model for shadow economy uses the same set of independent variables as did the model for the informal economy, i.e., informal labor share (LABINFORMSHR), corruption (GOVCORRUPT) and total R&D spending (RANDDEXP).

$$\begin{aligned} GDPInformalCalc_{r,t} &= \mathbf{gdpshadowcoeffintercept} + \mathbf{gdpshadowcoefflabinf} \\ &* LABINFORMSHR_{r,t-1} + \mathbf{gdpshadowcoeffcorrupt} \\ &* GOVCORRUPT_{r,t-1} + \mathbf{gdpshadowcoeffRandD} * RANDDEXP_{r,t-1} \end{aligned}$$

where r and t denote country/region and time, respectively,

In the above equation, the basic values of the parameters are:

$gdpshadowcoeffintercept = 36.2814$, **$gdpshadowcoefflabinf = 0.1019$** , **$gdpshadowcoeffcorrupt = -1.8849$** , and **$gdpshadowcoeffRandD = -2.9572$** . The negative value on the corruption term is potentially misleading because in Transparency International's measure, higher values are more transparent and lower ones are more corrupt; the equation term increases GDPInformal as corruption increases.

As discussed above for the informal economy, initial values and base year model calculations for the size of the shadow economy will typically differ. Shift factors are used to adjust computed values to those from data in the base year, and the shift factors converge to zero over time.

For countries where the data on GDP share of shadow economy is greater than the informal labor share, indicating a somewhat counterintuitive situation in which the shadow sector seems to have a higher productivity per person than the formal economy, we accept those values in the model base year, but converge shadow GDP share to the informal labor shares over a period of 25 years.

3.5.5 Forward Linkages From Informal Economy

We have structured the forward linkages so that the model user can drive them with either the size of the informal economy (GDPINFORMSHR) or the size of the shadow economy (GDP SHADOWSHR). The logic is the same in either case. In the discussion below, which is used will depend on the value of the shadow economy switch ($gdpshadowon$). If that parameter is 0, as it is in the model base case, the driver is the informal economy share; if it is 1, the driver is the shadow economy share.

The fundamental driver of changes in economic productivity and government finance is the excess or deficit in the informal GDP share (ExDefInfGDP), relative to the expected value.

We begin by elaborating the manner in which that excess or deficit affects total (or multifactor) productivity and then turn to the manner in which it affects tax rates and therefore government revenues and expenditures.

3.5.5.1 Forward linkages: productivity and GDP

The key to the productivity linkage is to recognize that an informal economy (or a shadow economy but to a much lesser degree) has the tendency, particularly in developing countries, to generate an official GDP that is lower than the potential GDP (PotentialGDP), because labor is employed with less capital and less efficiency in the informal sector. The magnitude of this productivity cost of informality (InfCost) can be illustrated by the fact that the data for some economies suggest that while 60 percent or more of labor can be informal, only 20 percent of GDP might be informal. This suggests that each unit of informal labor in such a country is used only about one-third as productively as is formal labor (we compute the ratio as the inverse, suggesting that each surplus unit of informal or shadow economy in this example could have a productivity cost of 3; we further bound that ratio so that it cannot exceed 3 or fall below 1).

If $gdpshadowon=0$ then

$$InfCost_{r,t} = \frac{LABINFORMSHR_{r,t}}{GDPINFORMSHR_{r,t}}$$

else

$$InfCost_{r,t} = \frac{LABINFORMSHR_{r,t}}{GDP SHADOWSHR_{r,t}}$$

It is important to emphasize, however, that the productivity and GDP growth rates in the basic IFs model already reflect the size of each country's informal GDP share and whatever drag or boost that share is giving to growth. Moreover, normal patterns of change in the informal GDP share (and we have seen that the share tends to decrease over the long-run with either advance in adult formal education or GDP per capita, are essentially already built into the long-term dynamics of productivity change in IFs, for instance the basic convergence or catch-up of productivity in developing countries with higher income countries. Thus, our attention here is not to the basic impact of informality on growth, but the incremental impact that deviations from the normal or expected size of the informal sector and the normal or expected change in that size might have. We can refer to such deviation as excess or deficit informal GDP share (ExDefInfGDP).

This point is so important that it bears rephrasing and repeating in only slightly different language. Why is this excess or deficit important? The dynamics of the IFs system's basic forecasting of economic productivity and GDP and of government revenue and expenditure patterns are all built on historical patterns and estimated functions that represent the normal processes within countries of gradual reduction of informal economy size with socioeconomic development. Thus, a *normal* decrease over time in the informal economy size should not be allowed to further affect the dynamics of productivity or government finance. Similarly, if a country has a persistent and stably larger or smaller informal economy than would be expected at its level of socio-economic development, with GDP per capita at PPP serving as a proxy for that—it should have normal productivity and finance dynamics. It will be countries that change the share of the informal economy relative to their expected values with development, and that therefore demonstrate changes in the excess or deficit of the informal GDP share (ExDefInfGDP) that will alter their expected development pattern.

The discussion above detailed the calculation of the forecast value of the actual informal share of GDP (GDPINFORMSHR), so to compute the excess or deficit informal share we need only to explain the calculation of the expected share (GDPInformShrExp). That calculation uses a function estimated against GDP per capita at PPP, namely "GDP/Capita (PPP 2011) Versus Informal Share of GDP (Most Recent) Log."

$$GDPInformShrExp_{r,t} = F(GDPPCP_{r,t})$$

$$GDPShadowShrExp_{r,t} = F(GDPPCP_{r,t})$$

This expected level is then subtracted from the actual forecast of the informal share, and the difference is divided by 100 to calculate the excess or deficit (ExDefInfGDP) in the extent of informal activities in ratio rather than percentage terms.

If gdpshadowon=0 then

$$ExDefInfGDP_{r,t} = (GDPINFORMSHR_{r,t} - GDPInformShrExp_{r,t})/100$$

else

$$ExDefInfGDP_{r,t} = (GDPShadowSHR_{r,t} - GDPShadowShrExp_{r,t})/100$$

What might be the potential GDP of an economy with an informal percentage share that is larger than expected? We can calculate that by multiplying the Official Economy (GDP) by the excess or deficit share (ExDefInfGDP) and by the unit informal cost (InfCost), adding that to base share of the official economy—the base share is the official economy minus that excess or deficit informal portion and should not be understood to be entirely formal, but rather to be the expected core mix of formal and informal.

$$\begin{aligned} PotentialGDP_{r,t} \\ = OfficialGDP_{r,t} * (1 - ExDefInfGDP_{r,t}) + OfficialGDP_{r,t} \\ * ExDefInfGDP_{r,t} * InfCost_{r,t} \end{aligned}$$

To illustrate this more concretely, if a country had an official economy of \$100 billion, but 20 percent of that were "excess" informal economy, the base economy would be \$80 billion. If the unit cost of excess informality were a factor of 3, the excess \$20 billion of informal economy could potentially become \$60 billion and boost the potential GDP to \$140 billion. Thus the calculation of the potential GDP in the formula above would be $100*(1-0.2) + 100* 0.2*3 = 80+60=140$.

For the calculation further below of impact on productivity of reducing (or increasing) this potential GDP, we need a ratio of the official to potential GDP and will need to look at that ratio over time. Getting that ratio only requires a re-ordering of the terms in the above equation:

$$\frac{OfficialGDP_{r,t}}{PotentialGDP_{r,t}} = \frac{1}{(1 - ExDefInfGDP_{r,t}) + ExDefInfGDP_{r,t} * InfCost_{r,t}}$$

We call the ratio of official to potential GDP the (potential) adjustment factor for the total factor productivity and thus economy (*TefInfAdjust*).

$$TefInfAdjust_{r,t} = \frac{OfficialGDP_{r,t}}{PotentialGDP_{r,t}}$$

Again, we can clarify by extending our concrete illustration. For the economy above the right hand side is $1/((1-0.2)+0.2*3)=1/(0.8+0.6)=1/1.4=0.714$. Assume that the country were able to reduce its excess informal economy to 10 percent. The equation would then become $1/((1-0.1)+0.1*3)=1/(0.9+.3)=1/1.2=0.833$.

If the country accomplished this reduction (which might take 10 or 20 years to do, of course), what would be the impact on economic production? Basically, it should be able to boost the overall magnitude of production by $0.833/0.714=1.167$, that is, by 17 percent.

The change in the official/potential ratio from year-to-year (*TefInfAdjustFac*) is what will affect productivity and economic growth. The model user can also modify the impact with an adjustment multiplier parameter (*tefinfadjm*).

$$TefInfAdjustFac_{r,t} = \frac{TefInfAdjust_t}{TefInfAdjust_{t-1}} * tefinfadjm_{r,t}$$

We then feed that change into a variable that the model maintains for total factor productivity (TEFF). That variable is a stock, so that it carries over from year to year and has been adjusted already by other drivers of productivity before it is adjusted below by the informal economy term. Thus, our example country making a 10 percent reduction in excess informal GDP over 10 years might actually make a 1.6 percent change for each of the 10 years. Those annual changes would cumulatively affect TEFF in such a way that after 10 years, the total impact should be roughly that of the 17 percent computed above

(roughly, because compounding and internal model dynamics would change that somewhat). $TEFF$ is maintained by economic sector (s).

$$TEFF_{r,t,s} = TEFF_{r,t,s} * (1 + TefInfAdjustFac_{r,t})$$

The variable $TEFF$ is a multiplicative term in the production function and the computation of value added and GDP. Thus, any change in $TEFF$ feeds through to GDP in the same proportion.

3.5.5.2 Forward linkages: government revenue

The general logic for computing the impact on productivity and GDP from changes in excess or deficit informal GDP share is also used for computing its impact on government revenues, except that there is no need for a unit cost term. Except for external items such as foreign aid, government revenues in IFs come from four taxes: those on firm profits ($FIRMTAX$) and on household income ($HHTAX$), plus specialized taxes such as indirect taxes on consumption ($INDIRECTTAX$) and contributions to social welfare programs or pensions ($SSWELTAX$). In each case, IFs also has a variable that specifies the rate of the tax on the appropriate base ($FIRMTAXR$, $HHTAXR$, $INDIRECTTAXR$, and $SSWELTAXR$).

The simplifying assumption we make is that informal economies contribute very little or nothing to the true base for these taxes. Thus the tax rates (coming to us along with tax bases and amounts from official data) are actually in the real world associated with the formal sub-portion of the official economy and appropriate tax bases in that official economy. As the share of the informal economy decreases, we can expect those bases to expand and to generate effectively higher tax rates. In IFs, we can use this general understanding to increase (or decrease) the tax rates accordingly as the excess or deficit informal economy variable decreases (or increases).³ We can compute a tax rate adjustment factor ($TaxInfAdjustmentFac$) that reflects this, also introducing a multiplicative parameter usable for scenario analyses (*taxinfadjm*).

$$\begin{aligned} TaxInfAdjustmentFac_{r,t} \\ = 1 + (ExDefInfGDP_{r,t=1} - ExDefInfGDP_{r,t}) * taxinfadjm_{r,t} \end{aligned}$$

Because the adjustment factor is computed relative to the first year (t=1), the change in all four tax rates should also be applied to the rate of the first year. (In the base of households there are two types (h), skilled and unskilled.)

$$FIRMTAXR_{r,t} = FIRMTAXR_{r,t=1} * TaxInfAdjustmentFac_{r,t}$$

$$HHTAXR_{r,h,t} = HHTAXR_{r,h,t=1} * TaxInfAdjustmentFac_{r,t}$$

³ The model computes $HHTax = HHIncOfficial * HHTaxR$, but the actual situation we are trying to represent is $HHTax = HHIncFormal * (HHTaxR * X)$ where X is an implicit factor of increase in the real tax rate so as to represent the application of it only to the formal economy. Some algebra would change the necessary adjustment in the tax rate, but it should not be highly significant relative to other sources of uncertainty.

$$INDIRECTTAXR_{r,t} = INDIRECTTAXR_{r,t=1} * TaxInfAdjustmentFac_{r,t}$$

$$SSWELTAXR_{r,t} = SSWELTAXR_{r,t=1} * TaxInfAdjustmentFac_{r,t}$$

Total government revenues are responsive primarily to these tax rates and their bases and the model includes dynamics that also link government expenditures to taxes and fiscal balances. Thus changes in these revenues will have pervasive effects through the rest of the model.

3.6 The Adjustment Mechanism and Other Specialized Functions

There are many specialized functions used in the IFs forecasting system. See the IFs Development Manual for more detail on them. Here we describe only the one that has special significance for the energy model.

In many models of IFs, especially the economic model and the two elaborated sectoral models (agriculture and energy), IFs relies upon an adjustment function to alter key variables (e.g., demand, prices, trade, and investment) in the pursuit of equilibrium. The adjustment function compares the level of some stock type variable (most often either inventory levels or prices, but including other variables such as international indebtedness) with a desired level, and adjusts the dependent variable.

IFs computes a difference (DIFF1) between the actual and desired levels and scales that difference with a scaling base (SCALINGBASE) value (for instance, total production in an economic sector might be a reasonable scaling base value against which to gauge the importance of a deviation of inventories from desired levels). In addition, the adjustment mechanism uses a second-order difference (DIFF2) to compare the level of the stock-type driving variable with its value in the previous time cycle, relying upon the same scaling base.

Non-zero differences result in a multiplier value (MUL) that deviates from "1" depending on the magnitude of two elasticities (EL1 and EL2). Specifically, the formulation is

$$MUL = \left(1 + \frac{DIFF1}{SCALINGBASE}\right)^{EL1} * \left(1 + \frac{DIFF2}{SCALINGBASE}\right)^{EL2}$$

This mechanism is represented in a function called the adjuster (ADJSTR) that the model calls at numerous locations. The magnitude of the two parameters will, of course, differ depending on the model variable in which equilibrium is being pursued. Experience has shown, however, that EL1 normally takes absolute values between 0.2 and 0.4, while EL2 is most often two times the value of EL1 and thus varies most often between 0.4 and 0.8. The values of EL1 and EL2 have been determined experimentally, in order to be large enough to maintain approximate equilibrium and small enough to avoid unreasonably rapid or extreme oscillation. There will inevitably be some oscillation in equilibrium-seeking processes, and in some cases (such as inventory levels), the values could be set so as to provide an oscillation consistent with known cycles (such as business cycles). Because IFs is a long-term rather than a short-term model, however, we

have generally devoted little attention in scaling to the oscillation cycle, focusing instead on long-term stability in the face of shocks introduced by scenarios of model users.

This kind of adjustment mechanism is sometimes called a PID controller, that is, an adjustment process that responds proportionately (the adjustment parameters) to the integral of the error (the stock discrepancy) and to the derivative of the error (the change in stock term).. For more information see the books by Chang (1961) and by Mishkin and Braun (1961). An early version of this adjustment mechanism was developed by Thomas Shook for the Mesarovic-Pestel modeling project.

3.7 Creating and Initializing Country SAMS: The Pre-processor

Preparing an initial data load for a model sometimes requires almost as much work as does creating and maintaining the dynamics of the model. Data inconsistency and data holes require attention; in a model like IFs with physical representations of partial equilibrium sectoral models (agriculture and energy) as well as a general equilibrium multi-sector model represented in value terms, there is also the need to reconcile the physical and value data.

Creation of a data pre-processor within IFs moved the project from manual handling of issues around data loads to automatic, algorithm-based processing. The pre-processor greatly facilitates both partial data updates as better data become available and rebasing of the entire model to a new initial year (such as the rebasing from 2005 to 2010). It works with an extensive raw data file for all areas of the model, using data gathered for 1960 through the most recent year available. This allows it to create an historical data load (based in 1960) for the purposes of historical validation analysis, as well as the load for forecasting.

It is not the purpose here to fully document the pre-processor, but a summary description is important (see Hughes and Irfan 2006 for much more detail). In general, the pre-processing begins with demographics, and imposes total population data on the cohort-specific data by normalizing cohort numbers to the total. The pre-processor reads values for a wide range of population-related variables: total fertility rate, life expectancy, crude birth rate, crude death rate, urban population, migration etc. IFs uses cross-sectionally estimated relationships to fill holes in such data (generally with separate functions for the 1960 (historical base) and 20xx (future) data loads). Most often, functions driven by GDP per capita at PPP have had the highest correlations with existing data; the best functions have often been logarithmic, because the most rapid structural change occurs at lower levels of GDP per capita (Chenery 1979; Hughes 2001). The philosophy in demographics and in subsequent issue areas in the pre-processor is that values for all countries in IFs will come from data when they are available, but will be estimated when they are not.

The pre-processor then proceeds to the agricultural and energy issue areas. In agriculture, the pre-processor reads data on production and trade. It aggregates production of various crops into a single crop production variable used by the model. It similarly aggregates meat and fish production for the model. It computes apparent consumption. It reads data on variables such as the use of water and on the use of grain for livestock feed. It uses

estimated functions or algorithms to fill holes and to check consistency (for instance, checking grain use against livestock herd and grazing land data).

In energy, the pre-processor reads and converts energy production and consumption to common units (billion barrels of oil equivalent). It checks production and reserve/resource data against each other and adjusts reserves and resources when they are inconsistent. Null/missing production values are often overridden with a very small non-zero value so that a “seed” exists in a production category for the subsequent dynamics of the model (a technique used by the Interfutures model of the OECD). World energy exports and imports are summed; world trade is set at the average of the two and country-specific levels are normalized to that average.

The outputs from processing of agricultural and energy data become inputs to the economic stage of pre-processing. The economic processing begins by reading GDP at both exchange rates and purchasing power and saving the ratio of the two for subsequent use in forecasting. The first real stages of economic data pre-processing center on trade. Total imports and exports for each country are read; world sums are computed and world trade is set at the average of imports and exports; country imports and exports are normalized to that global average. The physical units of agricultural and energy trade are read and converted to value terms. Data on materials, merchandise, service, and ICT trade are read. Merchandise trade is checked to assure that it exceeds food, energy, and materials trade, and manufactures trade is identified as the residual. All categories of trade are normalized. When this process is complete, the global trade system will be in balance. The use in IFs of pooled trade rather than bilateral trade makes this easier, but a similar process could be used for bilateral trade with Armington structures.

The processes for filling the SAM with goods and services production and consumption, and with financial flows among agent-classes follow next. More information can be found in the documentation of the economic model. The pre-processor then moves on to other models in IFs, including education, health, infrastructure and socio-political ones.

See Hughes and Hossain (2003) for earlier documentation on the SAM and also its initialization in the pre-processor. See Hughes with Irfan (2006) for documentation on the pre-processor.

4. References

- Allen, R. G. D. 1968. *Macro-Economic Theory: A Mathematical Treatment*. New York: St. Martin's Press.
- Barker, T.S. and A.W.A. Peterson, eds. 1987. *The Cambridge Multisectoral Dynamic Model of the British Economy*. Cambridge: Cambridge University Press.
- Barro, Robert J. 1997. *Determinants of Economic Growth: A Cross-Country Empirical Study*. Cambridge, Mass: MIT Press.
- Barro, Robert J. and Xavier Sala-i-Martin. 1999. *Economic Growth*. Cambridge, Mass: MIT Press.
- Brockmeier, Martina and Channing Arndt (presenter). 2002. Social Accounting Matrices. Powerpoint presentation on GTAP and SAMs (June 21). Found on the web.
- Bussolo, Maurizio, Mohamed Chemingui and David O'Connor. 2002. A Multi-Region Social Accounting Matrix (1995) and Regional Environmental General Equilibrium Model for India (REGEMI). Paris: OECD Development Centre (February). Available at www.oecd.org/dev/technics.
- Chen, Martha Alter. 2012. "The Informal Economy: Definitions, Theories, and Policies." WIEGO Working Paper no. 1. Women in Informal Employment Globalizing and Organizing (WIEGO), Cambridge, MA.
- Chenery, Hollis, and Moises Syrquin. 1975. *Patterns of Development, 1950-1970*. London: Oxford University Press for the World Bank.
- Chenery, Hollis. 1979. *Structural Change and Development Policy*. Baltimore: Johns Hopkins University Press.
- Dimaranan, Betina V. and Robert A. McDougall, eds. 2002. *Global Trade, Assistance, and Production: The GTAP 5 Data Base*. Center for Global Trade Analysis, Purdue University. Available at http://www.gtap.agecon.purdue.edu/databases/v5/v5_doco.asp.
- Duchin, Faye. 1998. *Structural Economics: Measuring Change in Technology, Lifestyles, and the Environment*. Washington, D.C.: Island Press.
- Globerman, Steven. 2000 (May). Linkages Between Technological Change and Productivity Growth. Industry Canada Research Publications Program: Occasional Paper 23.
- Griffith, Rachel, Stephen Redding, and John Van Reenen. 2000. *Mapping the Two Faces of R&D: Productivity Growth in a Panel of OECD Industries*. Institute for Fiscal Studies (January)

Hughes, Barry B. 1987. "Domestic Economic Processes," in Stuart A. Bremer, ed., *The Globus Model: Computer Simulation of Worldwide Political Economic Development* (Frankfurt and Boulder: Campus and Westview), pp. 39-158.

Hughes, Barry B. 2001. "Global Social Transformation: The Sweet Spot, the Steady Slog, and the Systemic Shift," *Economic Development and Cultural Change* 49, No. 2 (January): 423-458.

Hughes, Barry B. and Anwar Hossain. 2003. "Long-Term Socio-Economic Modeling: With Universal, Globally-Integrated Social Accounting Matrices (SAMs) in a General Equilibrium Model Structure." IFs Project Living Document, University of Denver.

Hughes, Barry B. 2005. "Forecasting Productivity and Growth with International Futures (IFs) Part 1: The Productivity Formulation," Pardee IFs Working Paper 2005.05.01, available at http://pardee.du.edu/sites/default/files/WP_2005_05_Productivity_Growth_IFs_Part1_0.pdf.

Hughes, Barry B. with Mohammad T. Irfan. 2006. "The Data Pre-Processor of International Futures (IFs). IFs Project Living Document, University of Denver.

Hughes, Barry B., Mohammad T. Irfan, Haider Khan, Krishna Kumar, Dale S. Rothman, and José Roberto Solórzano. 2009. *Reducing Global Poverty*. vol. 1, Patterns of Potential Human Progress series. Boulder, CO, and New Delhi, India: Paradigm Publishers and Oxford University Press.

International Labour Organization (ILO). 2013. *Measuring Informality: A Statistical Manual on the Informal Sector and Informal Employment*. Geneva: International Labour Office.

Jansen, Karel and Rob Vos, eds. 1997. *External Finance and Adjustment: Failure and Success in the Developing World*. London: Macmillan Press Ltd.

Kehoe, Timothy J. 1996. *Social Accounting Matrices and Applied General Equilibrium Models*. Federal Reserve Bank of Minneapolis, Working Paper 563.

Khan, Haider A. 1998. *Technology, Development and Democracy*. Northhampton, Mass: Edward Elgar Publishing Co.

Kornai, J. 1971. *Anti-Equilibrium*. Amsterdam: North Holland.

Kuznets, Simon. 1959. "On Comparative Study of Economic Structure and Growth of Nations." In *The Comparative Study of Economic Growth and Structure*. New York.

———. 1966. *Modern Economic Growth: Rate, Structure, and Spread*. New Haven: Yale University Press.

Lee, Ronald, Andrew Mason, and Timothy Miller. 1998 "Life Cycle Saving and the Demographic Transition in East Asia." Working Paper for *Population and Economic Change in East Asia* a supplement to Volume 26 of *Population and Development Review*.

Lee, Ronald D. and Andrew Mason, eds. 2011. *Population Aging and the Generational Economy: A Global Perspective*. Cheltenham, UK: Edward Elgar. Available at <http://idl-bnc.idrc.ca/dspace/bitstream/10625/47092/1/133467.pdf>.

McMahon, Walter W. 1997. *Education and Development: Measuring the Social Benefits*. Oxford: Oxford University Press.

Modigliani, Franco. 1976. "Life-cycle, individual thrift, and the wealth of nations," *American Economic Review*, 76(3), 297–313.

Narayanna, G., Badri Angel Aguilar, and Robert McDougal, eds. 2012. *Global Trade, Assistance, and Production: The GTAP 8 Data Base*. Center for Global Trade Analysis, Purdue University. Available at https://www.gtap.agecon.purdue.edu/databases/v8/v8_doco.asp

North, Douglass C., John Joseph Wallis, and Barry R. Weingast. 2009. *Violence and Social Orders: A Conceptual Framework for Interpreting Recorded Human History*. Cambridge: Cambridge University Press.

Organisation for Economic Co-operation and Development (OECD). 2002. *Measuring the Non-Observed Economy: A Handbook*. Paris: OECD Publishing. doi: 10.1787/9789264175358.

Pan, Xiaoming. 2000 (January). "Social and Ecological Accounting Matrix: an Empirical Study for China," paper submitted for the Thirteenth International Conference on Input-Output Techniques, Macerata, Italy, August 21-25, 2000.

Pesaran, M. Hashem and G. C. Harcourt. 1999. Life and Work of John Richard Nicholas Stone. Available at <http://www.econ.cam.ac.uk/faculty/pesaran/stone.pdf>.

Pistaferri, Luigi. 2001. "Superior Information, Income Shocks, and the Permanent Income Hypothesis," *The Review of Economics and Statistics* 82 (3), August: 465-476. Available at <http://web.stanford.edu/~pista/restat.pdf>.

Pyatt, G. and J.I. Round, eds. 1985. *Social Accounting Matrices: A Basis for Planning*. Washington, D.C.: The World Bank.

Romer, Paul M. 1994. "The Origins of Endogenous Growth," *Journal of Economic Perspectives* Vol 8, No. 1 (Winter): 3-22.

Sachs, Jeffrey. 2005. *The End of Poverty: Economic Possibilities for Our Time*. New York, NY: Penguin Press.

- Solow, Robert M. 1956. "A Contribution to the Theory of Economic Growth," *Quarterly Journal of Economics* 70, 1 (February): 65-94.
- Schneider, Friedrich and Dominik H. Enste. 2000. "Shadow Economies: Size, Causes, and Consequences." *Journal of Economic Literature* 38: 77—114. doi: 10.1257/jel.38.1.77
- Stone, Richard. 1986. "The Accounts of Society," *Journal of Applied Econometrics* 1, no. 1 (January): 5-28.
- Syrquin, Moshe, and Hollis Chenery. 1989. "Three Decades of Industrialization." *The World Bank Economic Review* 3 (2): 145–81.
- Taylor, Lance. 1979. *Macro Models for Developing Countries*. New York: McGraw-Hill.
- Thorbecke, Erik. 2001. "The Social Accounting Matrix: Deterministic or Stochastic Concept?", paper prepared for a conference in honor of Graham Pyatt's retirement, at the Institute of Social Studies, The Hague, Netherlands (November 29 and 30). Available at <http://people.cornell.edu/pages/et17/etpapers.html>.
- Zhang, Jie, and Junsen Zhang. 2005. "The Effect of Life Expectancy on Fertility, Saving, Schooling and Economic Growth: Theory and Evidence*." *Scandinavian Journal of Economics* 107(1): 45–66. doi: 10.1111/j.1467-9442.2005.00394.x.
- Zhang, Jie, and Junsen Zhang. (2009). "Longevity, Retirement, and Capital Accumulation in a Recursive Model with an Application to Mandatory Retirement." *Macroeconomic Dynamics* 13: 327-348.