

CGE Methods for Poverty Incidence Analysis: An Application to Vietnam's WTO Accession

David Roland-Holst[†]
Mills College and UC Berkeley

January, 2004

Abstract

Greater market access can be a potent catalyst for poverty alleviation in the rural communities of emerging economies. This is particularly true for more remote areas that lie outside the networks of transport and other infrastructure that link urban areas. These areas are also more isolated from administrative and informal linkages that could facilitate market access, extension service, and technology diffusion. Finally, lack of market participation compounds itself by limiting capital accumulation and the potential to invest for expanded production capacity. Likewise, external reform and greater outward orientation at the national level can be a catalyst for growth, leveraging external markets to expand productive capacity and resource use in the domestic economy. This paper examines the linkages between these two phenomena with a new approach to micro-macro CGE modelling.

The micro-simulation approach developed here can significantly improve visibility for policy makers who seek to strengthen the linkage between national growth policies and local income potential among the poorest rural households. Isolated and subsistence oriented small holders, particularly in mountainous regions, face significant barriers to market access and highly distorted prices. In this research, we seek to elucidate the microeconomic channels by which externally driven price reform can promote market participation, agricultural diversification, and transition from poverty to higher and more sustainable levels of income and savings.

1. Introduction

It has long been apparent that greater outward orientation and domestic market liberalization can stimulate economic growth. More detailed incidence of trade liberalization and market integration is much less obvious, however, yet this information is indispensable to policy makers who seek to propagate the gains from external reform and more effectively anticipate and deal with adjustment issues. Fortunately, a

[†] Paper presented at the Seventh Annual Conference on Global Economic Analysis: Trade, Poverty, and the Environment, June 17 - 19, The World Bank, Washington, D.C All opinions expressed here are those of the authors and should not be attributed to their affiliated institutions. Special thanks are due Saule Kazybayeva, Sarah Baird, and Tomoki Fujii for excellent research assistance, and to Maurizio Bussolo and Dominique van der Mensbrugghe for invaluable comments. Contact: dwrh@are.berkeley.edu

confluence of improved micro data, economic models, and computing resources promises better visibility on these essential aspects of the development process.

This paper reports on a World Bank project to develop new empirical tools for analysing rural income determination, applied to the context of WTO accession to a region of Vietnam with very high levels of poverty. In particular, we develop a new approach to micro-macro CGE modelling that better captures rural sector production/consumption and its linkages to regional, national, and international markets. This is done with a synthesis of the multi-market approach with economywide CGE techniques. In particular, we seek to extend micro-macro modelling to incorporate market segmentation and its attendant price dispersion. Previous efforts to exploit detailed household data in CGE models have relied on national market clearing and homogeneous prices in the underlying product, factor, and asset markets. By contrast, we believe price dispersion is a defining characteristic and determinant of inequality, for developing countries generally and the rural sector in particular. Indeed, these price differences embody essential information about the structural impediments to economic progress.

To advance research on the detailed incidence of national and sectoral policies on particular categories of households located in particular regions of a country, we propose an approach that integrates: (1) a consistent nationwide CGE model with (2) a linked regional multi-market model, and (3) micro-simulation of impacts at the household level based on individual records from the recently completed VHLSS. Components (1) and (3) reproduce the approach of established micro-macro CGE models. The integration of component (2) as the link between (1) and (3) is the main novelty of the proposed approach.

Markets can then close at different levels according to degree of market integration: at the household level when there is market failure (household non-tradables), at the regional level if goods are only traded in the region (regional non-tradables), at the level of the country when there is trade across regions (national non-tradables), and using border prices for tradable goods. Trade liberalization bears directly on tradable goods and indirectly on the others. Liberalization in services amounts to moving “one level up” in competition for non-tradables. The growth and income distribution effects of different schemes of market integration can be analyzed through reconfiguration of how markets close at different levels. The impact of reduction of transactions costs and increasing market integration can also be simulated at the household level, as previous non-tradables are now sold or purchased.

Preliminary results from the model strongly support the notion that price dispersion is an essential determinant of inequality and, in rural areas with limited market access, a factor that appears to reinforce poverty. Policy experiments with the model indicate that facilitation of market access and its attendant price convergence can be a potent catalyst for propagating trade-induced growth benefits and poverty alleviation.

2. Methodological Overview

What follows is a general and rather heuristic summary of the modelling approach. As the development stage progresses, these techniques will be more refined by

empirical testing and specification choices informed by econometric analysis. When the prototype model is completed and implemented for the project research, it will be documented in full technical detail.

2.1. Analytical Approach

After summarizing three existing standards for this type of analysis, we propose a new approach that we believe can advance practical understanding of trade and growth linkages at the more microeconomic level. Trade and aggregate growth may be highly correlated, but poverty is ultimately a microeconomic experience and one that can persist and even intensify in a growing economy. To elucidate the incidence of growth effects more completely, a variety of multi-sectoral models has been developed.

1. The traditional, national level CGE approach - This consists in economy-wide CGE models calibrated to Social Accounting Matrices (SAMs) and other structurally detailed national data. These models offer more detail than macro models, but usually specify national market interactions between representative consumers and sectoral representative producers. Because of this, they really capture only mean/median behavior of households and firms and have relatively little to say about functional or spatial economic diversity (both of which are very important to real incidence in developing countries). Especially in the rural sector, this aggregative approach to agent and market modelling is empirically unsatisfactory.

2. "Micro-Macro" CGE models¹ - These models take advantage of the advent of detailed surveys to disaggregate the household side of the economy. In most applications to date, this entailed modelling with national markets and representative firms as usual, but tracing income effects from (national) price changes to individual household assets, consumption patterns, and (in some cases) factor supply. While the approach used new data to provide information about portfolio effects, it still relies on homogeneous economywide markets and traditional production structures. The latter feature also retains its deficiencies for modelling of the rural sector, where household production units behave very differently from nationally representative sectoral production functions.

3. "Village" CGE Modelling (e.g. Taylor and Adelman: 1992) - In this approach, a CGE model is calibrated for a localized community economy and this economy is embedded in national accounts or even in a national CGE model. Some work has been done to string together a mosaic of such models, but this has yet to extend beyond sub-regional economies, and it is very unlikely that national policy will be supported by an inclusive constellation of such models in the foreseeable future.

4. Integrative micro-macro modeling – Each of the above approaches adds something to our empirical understanding of market-mediated behavior, but each still has significant limitations for application in the present context. The national model approach is consistent, but lacks detailed incidence information that is highly desirable for policy makers. The prevailing standard for micro-macro simulation sharpens this perspective in

¹ This approach was pioneered by Bourguignon and a variety of co-authors. See e.g. Atalas and Bourguignon (2000) and more recently Robilliard et al (2001), references in this overview are indicative rather than exhaustive.

the right direction by linking national results to local outcomes, but it has been extensively criticized for lack of endogeneity or consistent synthesis (micro-macro closure). The village approach is interesting in its own right, but again lacks any aggregate consistency framework. To advance research on the detailed incidence of national and sectoral policies on particular categories of households located in particular regions of a country, we propose a fourth approach that extends existing work in three ways:

1. We attempt a more integrated approach to micro-macro modelling, proposing a data stratification strategy that enables synthesis of both perspectives into a single numerical solution process.
2. Households in the agricultural sector will be modelled as more integrated production/consumption units.
3. Explicit account will be taken of price dispersion. While traditional models assume national market clearing and nationwide homogeneous prices, we calibrate prices to local conditions in an effort to capture price dispersion and its affects on household and enterprise balance sheets and incentives.

Each of these extensions contributes new insights from empirical modelling of rural sector households, and in the next section we describe how these innovations will be implemented. We discuss this approach in greater detail in Section 2 below.

2.2. Data Resources

For the proposed Northern Mountain Region, we are fortunate that the data for this exercise are already extant. These include a new and detailed national SAM, a regional SAM, and detailed (VHLSS) household surveys. The national SAM was developed by Roland-Holst, Tarp, and colleagues under Danida sponsorship (Rand et al: 2002) and the NMR SAM was developed by Roland-Holst last year under JBIC sponsorship. Moreover, these data already detail the composition of farm production at all three levels of aggregation (national, regional, and household). Roland-Holst and Tarp are also updating the 2000 national SAM this summer to take account of a new input-output table and improved trade data.

3. Methodology

The starting point for our analytical extensions consists of three components, a conventional economywide CGE model, a micro simulation module derived from Robilliard et al (2001), and a database consisting of two social accounting matrices and several household surveys. In this section, we discuss the first two, and how they will be extended and synthesized into a more integrated framework. Throughout the model

development process, specification choice will be informed by econometric analysis of Vietnamese micro data.

The base CGE is a prototype model for Vietnam developed a few years ago by Roland-Holst and van der Mensbrugghe.² This represents a state-of-the-art but standard dynamic CGE model of the national economy, including homogeneous national markets/prices, neoclassical production sectors, and very aggregate representative households. In the discussion that follows, we describe how this prototype will be modified to incorporate the desired extensions. As all the discussion applies equally to a model at the national or a single province (e.g. the Northern Uplands) level, we make no further reference to this distinction for the moment.

3.1. Endogenous Treatment of Macro and Micro Behavior

For the specification of macroeconomic variables, we rely on the standard CGE model, consistently aggregating demand, supply, and other institutional activities into accounts consistent with NIPA and other macro balances. For more detail on this, the reader is referred to the Vietnam prototype model documentation in Roland-Holst and van der Mensbrugghe (2002).

It has already been emphasized, however, that lack of microeconomic detail is a major shortcoming of such models. In response to this, several approaches have emerged, the main ones being summarized in the previous section. The current standard for this kind of work is the so-called micro-macro approach developed by Alatas and Bourguignon (2000) and recapitulated in Robilliard et al (2001). These authors seek to elucidate the detailed effects of national policies by linking two separate simulation frameworks, a relatively standard national CGE and a micro simulation framework calibrated to some LSMS survey. The two modules are linked together in a one-time or iterative sequence, using national level results as inputs to the micro simulation. Generally speaking, national prices of goods and factors are used to compute detailed household incomes, asset values, and some factor and output supply responses.

The advantage of this approach is of course the richness of detail for incidence results. It's most commonly discussed weakness is segmentation of the two perspectives, specifying the micro and macro behavior in different models with limited feedback and nothing that could really be called overall analytical regularity. Regular general equilibrium specifications of the national model are "linked" to a micro level blend of balance sheets, discrete choice econometric models, and partial equilibrium behavioural equations, with or without convergence criteria to yield some imputation of local effects arising from national markets. While all the components of these models are interesting and innovative, it is far from clear how they correspond to an integrative view of economic linkages. This ambiguity makes interpretation of closure rules and other systemic features such as price homogeneity and even uniqueness very difficult.

² This model is fully documented elsewhere and will not be reproduced here. See Roland-Holst et al (2002).

It has been argued that this two-part approach is necessary because incorporating fully disaggregated household survey data into an economywide model is unwieldy and unnecessary. We agree with both these arguments, but propose a different solution. Even with today's advanced computing resources, a CGE model with as many households as a typical LSMS would be awkward to implement. We also believe this may not be necessary, but for a different reason than would justify a companion micro-model.

In particular, we question uncritical acceptance of an existing LSMS sample as the standard for disaggregation in micro-level simulation modelling.³ The LSMS sample is itself a representative abstraction of a larger underlying population. Instead of running two models at radically different aggregation levels, we believe one model should be developed at an aggregation motivated by statistical analysis of the underlying sample. Given the limited number of economic characteristics usually considered in policy simulation exercises, it is reasonable to expect that some judiciously chosen sub-sample can represent the LSMS sample, just as the LSMS represents the population for a larger number of (survey) characteristics.

As an alternative to LSMS disaggregation as the standard for micro simulation, we propose econometric analysis to identify a more limited representative sub-sample that, with a defined set of explanatory variables (or a vector of attributes a_i), can explain variation in economic outcomes for the full sample (and by extension the population). What we are really proposing is a partial CGE disaggregation strategy based on variance criteria.⁴ An analogous technique has been recently applied to poverty mapping, where the goal was to extend an LSMS sample estimates to an underlying census population. Here we plan to develop a representative LSMS sub-sample, incorporate it directly into the CGE model, and then impute the CGE results back to the LSMS (and beyond).⁵

The main advantage of the proposed approach is to synthesize micro and macro simulation within a single modelling framework. This necessarily implies greater parsimony on the micro side and more detail on the macro side, but it is reasonable to expect a balance between the two to facilitate empirical analysis if it can be justified on statistical grounds. To see how the general procedure would work, consider a given LSMS household survey S with observations s_i , and a vector of attributes x_i drawn for each household. Consider also a vector y_i of economic variables relevant to each household, observable in the sample, that correspond to endogenous variables of an economywide model ("shared" variables in the micro-macro literature). There are then three stages to the proposed strategy:

³ This is not the main objection to the existing standard for micro-macro modeling, which has to do with ad hoc parameter adjustment and closure inconsistencies. We are also concerned about these, but the new approach obviates these issues.

⁴ This technique is based on small area statistics (e.g. Ghosh and Rao (1994), Rao (1999)) and econometric thinking pioneered by many authors including Hall (1992), Pakes and Pollard (1989), Hellerstein and Imbens (1999), and Keyser (2000). A seminal series of related empirical work has come from Elbers, Lanjouw, and Lanjouw (2000-2003).

⁵ Residuals that arise in this imputation back to the VLSS will not be removed by ad hoc parameter adjustment, but will instead be reported. With the right sub-sample stratification, these can be minimized.

1. Estimate generalized forms $y(x)$ using the LSMS sample, following a bootstrap procedure to derive a minimal representative sub-sample that explains a critical percent (e.g. > 90%) of the variance across the whole sample.⁶
2. Aggregate the LSMS and disaggregate the national level data to conform to representation by the sub-sample. Run the CGE calibrated to this database and obtain simulation results.
3. Impute the CGE results directly to the LSMS sample, and ultimately to the population, using the sampling information from Step 1.

The particulars of this approach will vary from application to application, but the general strategy can probably be applied to most LSMS samples. In the current project, we will experiment with several of variable pools and critical levels to ascertain the robustness of this approach. Analogous applications (e.g. Elbers et al:2003 on poverty mapping), indicate that hundredfold reductions in sample size have negligible effects on explained variation in target variables like poverty indexes.

Obviously, the resulting CGE model must also achieve synthesis of behavioural perspectives between the previous component approaches. The main implication here is more extended treatment of household behavior, including (at a minimum) detailed expenditure, factor supply, and production decisions. We elaborate on this issue next.

3.2. Modelling Rural Household Behavior

One of the most common complaints about applying CGE modelling to developing countries is that agricultural production bears little resemblance to a neoclassical/leontief specification of production. In this framework, a single representative production activity (for each agricultural product or sector) combines constant share inputs with factors of production recruited from national labor, capital and (sometimes) land markets. This activity then maximizes profits and some portion of its factor income accrues to rural households who make their expenditure decisions independently.

In reality, most rural sector households in the developing world (and over 50% in Vietnam) are combined production and consumption units. Their factors largely belong to them and a significant part of their output is retained for own consumption. Although evidence indicates that even remote households are aware of local market signals, more remote areas exhibit higher levels of self-sufficiency. This may be partly because of price biases such as those described above, or because of other behavioural characteristics. The consequence in any case is that more careful specification of rural household behavior is needed to capture this behavior.

While we make no pretence to definitive treatment of this, we believe that joint specification of production and consumption can shed important light on rural economic

⁶ See, e.g. Hall and Horowitz (1995) for background.

participation, particularly household decisions in response to changing prices and emergent market opportunities. During the model development process, we are experimenting with a variety of specifications, centered on a few general features.

Households are assumed to be characterized by the representative attribute vectors x_i drawn from the LSMS. This means the sample is stratified into household categories representing a discrete partition of each component in x_i as defined in the previous section. With the characteristics come econometrically estimated functions for significant behavioural relationships, including consumption, household production, and labor supply. The first two are standard, consumption calibrated to an AIDS or ELES form and production arising from a household profit function. For calibrating consumption, we assume households divide output between own use and the market. To calibrate production, we are fortunate to have estimates of both types of household production in the SAM. Activities with no own production in the base data will be assumed to be produced by standard neoclassical enterprises.

In the case of labor supply, the traditional micro-macro approach specifies discrete occupational choice calibrated to econometric qualitative choice models. This framework might be intellectually more satisfactory for individual labor supply, but it leads to many analytical problems with model closure and solution in either the two-part or synthetic model approach. By contrast, we believe our representative household subsample justifies use of a continuous specification, a CET function $lh_i(x_i, y)$ that yields a three-dimensional vector of household employment shares, divided between own employment, outside employment, and no employment. Such a functional form would be easily calibrated to VLSS shares, aggregate employment, and is smoothly responsive to relative wages.

3.3. Price Dispersion

The assumption of national market clearing and homogeneous prices is shared by both traditional national CGE models and more recent micro-macro models. Despite the wide use of this standard, however, price dispersion is a pervasive and robust characteristic of developing economies. Previous efforts to exploit detailed household data in CGE models have relied on national market clearing and implicitly homogeneous prices in the underlying product, factor, and asset markets.

Indeed, we would argue that, for developing countries generally and the rural sector in particular, price dispersion is a defining characteristic and determinant of inequality. It is often argued, for example that developing economies are replete with market failures. This may be true, but it does not mean that prices are failing to embody information about economic structure and conduct. Indeed, we believe that local price differences embody essential information about the impediments to economic progress.

Before presenting the formal treatment, it may be useful to motivate things with an example. Consider the agricultural terms of trade ($agtot$), localized to the individual household or village level. Depending upon distance and other barriers between this rural economic unit and a regional trading center, the $agtot$ can vary substantially. In particular,

greater remoteness can be expected to adversely affect agtot (from a rural perspective) in two ways, reducing the denominator (things rural households can sell) and increasing the denominator (things rural households must buy from outside). Both these components undermine rural purchasing power and incentives for market participation, and contribute directly to rural marginalization, poverty, and inequality. Over the longer term, upward bias in the prices of urban or international agricultural inputs, including agrochemicals and technology, retard adoption and agricultural productivity growth.

The mechanisms at work here also have implications for to price transmission. Simply put, more remote areas are less sensitive to border price variation and border markets are less responsive to market shocks in remote areas. The former implies, for example, that remote farmers will be less responsive to any change in border price incentives. This not only retards aggregate agricultural price response arising from international or metropolitan demand, but does so in a way that is biased geographically and probably regressive in income terms.

By symmetric logic, border or urban markets will be less responsive to rural price changes and the implied the needs of remote areas. This means that rural scarcity can only be alleviated by with greater local price increases, further undermining the purchasing power of rural households and increasing local financial risks arising from shortages.

In order to more fully capture these complex price and incentive characteristics, two approaches can be taken: market segmentation and calibrated margins. Both these phenomena are to some extent at work in any economy, especially one with significant structural impediments to market development. From a modelling perspective, however, the two are treated quite differently, so we need from the outset to determine which is preferable.

Using price and location data in the two VLSS surveys, we searched intensively for patterns of price correlation that might delineate segmented markets. Certainly, casual observation suggests that agents in rural areas habituate their trading within economic “watersheds” circumscribed by natural barriers, roadways, and other features that strongly influence transactions costs. Despite our efforts to identify the boundaries of such markets, however, we were unable to ascertain any systematic patterns of regionalized price determination. For this reason, we have decided to implement a specification of calibrated margins to explain price dispersion across the economy.

To accomplish this, the standard CGE model has been extended to incorporate four price wedges that reflect distribution margins:

$m_{x_{ki}}$ = margin between producer i price and the national market price

$m_{m_{ki}}$ = margin between import i border price and the national market price

$m_{e_{ki}}$ = margin between producer i and the border price

$m_{c_{ki}}$ = margin between consumer i price and the national market price

for any commodity k . In most models where margins are used, they are calibrated from initial data and held constant in *ad valorem* terms. In the present analysis, each of these margin rates will actually be endogenous variables, determined from a combination of

agent-specific characteristics and econometrically estimated parameters. For example, consider consumer i and commodity k . Then the model determines the corresponding consumption price as a function of

$$pc_{ki} = (1 + mc_{ki}(x_i))pd_{ki}$$

where pd_{ki} denotes the corresponding national market (CGE) price. This example shows margins depending on the vector of household characteristics x_i determined from the representative stratification of the VLSS, but analogous attributes are taken into account for the other margins (e.g. producers, import origins, and export destinations). Base margin rates will be calibrated to the initial data, as usual for tradable goods and border prices, and by reference to the micro data for consumer and producer prices.

Because price dispersion is endogenously associated with different agents, their behavior in response to differing incentives is captured in ways that would not be possible with a single, national price system. At the same time, the explicit inclusion of these margins allows a new genre of scenario analysis, namely policies to facilitate market access and openness.

3.4. Market Closure and Related Issues

Overall closure issues can exert important influences on economywide simulation results, and it is worth digressing on this issue a little to clarify our approach. We had originally proposed an approach that integrated:

- (1) a consistent nationwide CGE model with
- (2) a linked regional multi-market model
- (3) micro-simulation of impacts at the household level based on individual records from the recently completed VHLSS

Components (1) and (3) resemble the approach of established micro-macro CGE models. The integration of component (2) as the link between (1) and (3) was the main novelty of the proposed approach.

The original proposal envisioned segmented regional markets, as reflected in item (2). After intensive data analysis, however, we have concluded that (for Vietnam at least) detailed price data do not support any clear delineation of separate markets. Instead, we have concluded that price transmission behaves as though a national market were at work, but prices levels exhibit locational differences that more correctly reflect distribution margins. For this reason, we have chosen to adopt the margin calibration approach outlined in Section 2.3 above. This technique will achieve the same kind of synthesis (of items 1 and 3) we sought under a segmented market specification, but does so with more parsimony and transparent closure properties.

In the stratified price system we have specified, trade liberalization bears directly on tradable goods and indirectly on the others. Liberalization in services, for example, amounts to moving “one level up” in competition for non-tradables. The growth and income distribution effects of different market integration schemes can be analyzed

through varying the degree of pass-through implicit in the detailed distribution margins (elasticities in the functional forms for margins). The impact of reduction of transactions costs and increasing market integration can also be directly simulated at the household and producer level.

3. General Equilibrium Distribution and Incidence Measures

Significant progress has been made in recent years on methods for ex post analysis of income distribution and poverty. The advent of decomposable measures has inspired an extensive analytical literature and a broad spectrum of practical measurement tools.⁷ In this project, we exploit that literature and link it explicitly to the modeling framework, enabling a wide range of scenario analysis with endogenously determined distribution and incidence measures. There is precedent for this approach in more aggregated context, but the present effort seeks to capture more detailed institutional characteristics and market segmentation.⁸

In the present exercise, we want to elucidate incidence by exploiting new work on measures of poverty and income distribution. At the same time, we hope to extend these contributions by incorporating the novel features of our own modeling approach. In particular, we will present a set of incidence measures that explicitly incorporate local price dispersion, including some that take explicit account of rural household's threefold role as consumer, producer, and supplier of factor (mainly labor) services. Moreover, we will make these measures endogenous to the model, allowing them to vary directly from scenario to scenario rather than post-processing them.

By applying established measures to new (2002) Vietnamese panel data, we hope to improve general understanding about changing patterns of incidence in this country. By extending these measures to encompass more structural characteristics, we want to advance the same objective but also contribute to poverty research generally.

4. Data Resources

We are fortunate in the present case to have extensive and timely data, both at the micro and macro levels. The primary components of our data resources are a new Social Accounting Matrix for Vietnam, a regional SAM for the Northern Uplands, and several detailed household surveys. In this section, we provide a cursory survey of each.⁹

4.1. Social Accounting Matrices

The latest SAM for Vietnam, just completed for the year 2000, is the result of a three-year project to assemble and reconcile a variety of economywide data into a consistent set of tabular accounts. Generally speaking, the SAM provides a closed form, economywide accounting of linkages between activities (and/or commodities), factors,

⁷ See e.g. Bourguignon:1979, Foster et al:1984, and Kanbur:1984, followed by many contributions including Ravallion (examples follow) and others.

⁸ See e.g. Datt et al (2003).

⁹ See the references for full documentation of all these sources.

households, domestic institutions (e.g., investment, government), and foreign institutions in a tabular format that is transparent and amenable both to multiplier analysis similar to that popularized by Leontief and more sophisticated CGE analyses. These include for example studies focusing on the economic impact of initiatives such as WTO accession. The relevance of such analyses in the present phase of Vietnamese integration into the global economy and the international institutions dealing with trade issues can hardly be exaggerated.

An interesting SAM for Vietnam was published by the United Nations in the mid-1990s. While much of the theoretical analysis and overview in that document remains valid, that SAM is very aggregated and relies on a now outdated 1989 10-sector I/O table. Other contributions to this area include a SAM for Central Vietnam, underpinning the study by Bautista (2000), a working paper by Nielsen (2001) that estimated a 1997 SAM (included in the GTAP database), and the thesis by Huong (2000). However, until now there is no SAM reflecting the economic structure of Vietnam in the aftermath of the Asian financial crisis, taking into account the substantial changes in exports and imports over the past few years.

The first SAM in the current project, for 1999, was produced in 2002, and this has been followed by two 2000 tables, one prototype produced last year and a final one in February, 2004. The latter includes the new GSO input-output table for 2000, and represents the most up-to-date economywide data available from official sources, reconciled for the first time. The 2000 table is also the first SAM for Vietnam (or any other country as far as we know) with household and factor accounts calibrated directly to LSMS micro data. The general structure of the current SAM is summarized in the next table.

Vietnam SAM for 2000 – Structural Characteristics

1. Incorporates the new 2000 GSO Input-Output Table
2. 112 domestic production activities
3. 114 commodities (includes trade and transport margins)
4. 14 factors of production
5. 12 labor categories
6. Capital
7. Land
8. 16 household types, sampled from the VLSS
(Farmer, SelfEmp, Worker, UnEmp)x(Rural, Urban)xGender
9. 3 enterprises (Private, Public, and Foreign)
10. State (detailed fiscal instruments)
11. Consolidated capital account
12. 194 international trading partners

At a more detailed level, we also have a SAM that tabulates income and expenditure flows within the fourteen provinces of the target area of the present study, Vietnam's Northern Uplands. This table was estimated for 2000 under JBIC sponsorship, and is being updated to be conformal and otherwise consistent with the new national SAM.¹⁰

4.2. Micro Survey Data

Three LSMS type surveys for Vietnam provide detailed and direct observations on many economic and demographic characteristics of rural households, including information on household and community composition, asset/factor ownership, production and consumption patterns, local price information, immigration status, and many others.

To better understand what kinds of empirical analysis the household survey data will support, we have begun the project with extensive and intensive econometric analysis of the Vietnam Household Living Standards Survey (VHLSS). This includes autocorrelation studies of locational price dispersion to delineate market segmentation across the region, regression estimation of household expenditure, factor supply, and production functions for CGE model implementation, and more general correlation analysis and data mining to assess other sample features. Among the latter is the representative sub-sample strategy outlined in Section 2.

In these ways, direct analysis of the survey data has provided a variety of calibration resources for the CGE model, but may also produce results of independent interest, in which case they will be reported as part of the research documentation.

¹⁰ See IFPRI (2003) for more complete documentation.

Vietnamese Household Surveys used for this Study

<i>Name</i>	<i>Period of data collection</i>	<i>Sample size</i>	<i>Lowest level of representation</i>	<i>Types of income data collected</i>
1993 Vietnam Living Standards Survey	1992-1993	4800	Seven regions	Crop area, production, and sales; crop by-products; livestock income, fisheries income; livestock and fisheries expenses; forestry income; non-farm enterprises revenue and costs; wage income most important current wage jobs and most important wage jobs over past 12 months; remittances; land rent; social security; other income.
1998 Vietnam Living Standards Survey	1997-1998	6000	Ten strata (7 regions and 3 types of urban areas)	Almost identical content and structure as the 1993 VLSS.
2001-02 Vietnam Household Living Standards Survey	2001-2002	75,000	61 provinces	Similar to VLSS, but less detailed, particularly the sections on wage income and enterprise income.

5. The micro-macro modeling framework

5.1. The micro-simulation model

In the micro-simulation, we model the household income generation process.¹¹ Individuals make occupational choices and earn wages or profits accordingly. These labor market incomes plus exogenous other incomes, such as transfers and imputed housing rents, comprise household income. The micro-simulation enables us to take individual and household heterogeneity into account. Individual heterogeneity refers to personal characteristics, which influence occupational choices and income generated on the labor market. Occupational choices are subject to a number of factors, which include gender, marital status, or age of children. Important determinants of labor income are education and experience. Household heterogeneity is reflected, for example, in different sources of income and demographic composition. Furthermore, the micro-simulation captures some household heterogeneity in terms of expenditure structure. The micro-simulation is based on Vietnamese household surveys.¹²

Income Generation Model

The components of the income generation model are an occupational choice and an earnings model. Individual agents can choose between inactivity, wage-employment, and self-employment. In rural areas, there is a fourth option of being both wage-employed and self-employed. The occupational choice model is assumed to be slightly different for household heads, spouses, and other family members. As the possible occupational choices imply, earnings are generated either in the form of wages for employees or as profits for the self-employed. Individuals in rural areas can receive a mixed income from both types of activities. This latter option will be ignored in the following illustration of the model. Being self-employed means being part of what might be called a “household-enterprise”. All self-employed members of a household pool their incomes. This pooled income is then called profit. The mechanisms of profits earned in agriculture on the one hand side and other activities, such as petty trade, on the other are assumed to be different. Since agriculture plays a negligible role in urban areas, this differentiation is only implemented for rural areas.

The wage-employment market is segmented: the wage setting mechanisms are assumed to differ between urban and rural areas, for skilled and unskilled labor, and for females and males, which implies that there are eight wage labor market segments.

Household income comprises the labor income of all active household members and other income. Wages and profits are thus the endogenous income sources of the household. All other incomes are assumed to be exogenous and constant over time. The resulting total household income is deflated with a household group specific price index, which takes into account the differences in budget shares for food and non-food.

¹¹ The following section borrows heavily from Alatas and Bourguignon (2000), Robilliard et al. (2002) and others.

¹² The household survey used for estimation of the micro-simulation parameters is the 2002 Vietnam Household Living Standards Survey.

The income generation process, which consists of the occupational choice and the earnings models, is first estimated using data from the Vietnamese household survey from 2002.¹³ The estimated benchmark coefficients are then employed and changed in the micro-simulation.

Links to the CGE model

The micro-simulation and the CGE models are linked sequentially by a set of aggregate variables. Specifically, *firstly* the CGE calculates the new equilibrium for a specific scenario, and determines the following aggregate results: the average wage in each labor market segment, the average profits for different activities, the shares of self- and wage-employed for each segment (labor force composition), and the relative price of food and non-food commodities. *Then*, these aggregate variables are used as targets for the micro-simulation model where individual changes in earnings, labor force composition, and rural household output are computed. These micro changes are obtained by varying coefficients in the occupational choice, production, and earnings models. Coefficients are adjusted, and occupational choices, rural output, and earnings change accordingly, until the results of the micro-simulation are consistent, at an aggregate level, with the results from the CGE model.

Elements of the Model

The following set of equations describes the model. Household m has k_m members, which are indexed by i .e.

$$\log w_{mi} = a_{g(mi)} + x_{mi}\beta_{g(mi)} + e_{mi} \quad (1)$$

$$\log \pi_m = b_{f(m)} + z_m \delta_{f(m)} + F_{f(m)}(N_m, K_m) + \varepsilon_m \quad (2)$$

$$Y_m = \frac{1}{P_m} \left(\sum_{i=1}^{k_m} w_{mi} IW_{mi} + \pi_m \text{Ind}(N_m > 0) + y_0 \right) \quad (3)$$

$$P_m = s_{d(m)} p_f + (1 - s_{d(m)}) p_{nf} \quad (4)$$

$$IW_{mi} = \text{Ind} \left[c_{h(mi)}^w + z_{mi} \alpha_{h(mi)}^w + u_{mi}^w > \text{Sup} \left(0, c_{h(mi)}^s + z_{mi} \alpha_{h(mi)}^s + u_{mi}^s \right) \right] \quad (5)$$

$$N_m = \sum_{i=1}^{k_m} \text{Ind} \left[c_{h(mi)}^s + z_{mi} \alpha_{h(mi)}^s + u_{mi}^s > \text{Sup} \left(0, c_{h(mi)}^w + z_{mi} \alpha_{h(mi)}^w + u_{mi}^w \right) \right] \quad (6)$$

The first equation is a Mincerian wage equation, where the log wage of member i of household m depends on his/her personal characteristics. The explanatory variables include schooling years, experience, the squared terms of these two variables, and a set of regional dummies. This wage equation is estimated for each of the eight labor market segments. The index function $g(mi)$ assigns individual i in household m to a specific

¹³ The occupational choice model was estimated using a multinomial logit. The wage equations were estimated by Ordinary Least Squares. Correcting for selection bias in these equations did not lead to major changes in the results and was hence dropped. In the estimation of the profit functions, the number of self-employed was instrumented. For a more detailed discussion of the estimation methods see Alatas and Bourguignon (2000).

labor market segment. The residual term e_{mi} describes unobserved earnings determinants.¹⁴

The second equation represents the profit function of household m . Profits are earned if at least one member of the household is self-employed. The profit function consists of two components. The first is of a Mincer type human capital specification, including as explanatory variables the schooling of the household head, her/his experience plus the squared terms for the former two variables, and regional dummies. Secondly, we estimate profit arising directly from household production, which is of special relevance in the rural sector. Household output $F()$ depends on the number of self-employed in household m , N_m , as well as other non-human inputs or factors K_m . Finally, the residual ε_m captures unobserved effects. The index function $f(m)$ denotes whether a household earns profits in urban or rural areas. Different profit functions for agricultural, non-agricultural, and mixed activities are estimated in rural areas, with special attention to household production in the rural sector (discussed in greater detail below).

Family income is defined by the third equation. It consists of the wages and profits earned by the family members and an exogenous income y_{0m} . This exogenous income corresponds to “other income” in the survey and may include government transfers, transfers from abroad, capital income, etc.. IW_{mi} is a dummy variable that equals 1 if member i of the household is wage-employed and 0 otherwise. Likewise, profits will only be earned if at least one family member is self-employed ($N_m > 0$). Family income is deflated by a household specific price index.

This household specific price index is defined by equation (4). The parameter s denotes the expenditure shares for food- and non-food. These shares are calculated by household income quintiles. Note that the prices p_f for food and p_{nf} for non-food are generated in the CGE model. The index function $d(m)$ indicates to which of the five income brackets household m belongs and which food expenditure share is assigned to the household.

The fifth equation explains the aforementioned dummy IW_{mi} . The individual will be wage-employed if the utility associated with wage-employment is higher than the utility of being self-employed or inactive. The utility of being inactive is arbitrarily set to zero, whereas the utilities of the employment options depend on a set of personal and family characteristics, z_{mi} . These characteristics include gender, marital status, education, experience, other income, the educational attainments of other family members, and the number of children. Unobserved determinants of occupational choices are represented by the residuals.

Equation (6) gives the number of self-employed. Similar to the choice in equation (5), the individual i of household m will prefer self-employment if the associated utility is higher than the utility of inactivity or wage-employment. The self-employed household members form the “household enterprise” with N_m working members. Thus, the last two equations represent the occupational choices of the household members. The

¹⁴ It is important to note that the micro-simulation as specified here does not generate a synthetic panel. It rather produces a second cross-section. As will be explained later in more detail, we need to differentiate between permanent and transitory components of the residual in order to analyse income mobility or poverty transitions.

occupational choice model is estimated separately for household heads, spouses, and other family members in urban and rural areas. The index function $h(mi)$ assigns the individual to the corresponding group.

The model just described gives the household income as a non-linear function of individual and household characteristics, unobserved characteristics, and the household budget shares. This function depends on three sets of parameters, which are estimated based on the 2002 survey. These parameters include (1) the parameters of the wage equation for each labor market segment, (2) the parameters of the profit and production functions for “household enterprises” in urban areas and different activities in rural areas, (3) the parameters in the utility associated with different occupational choices for heads, spouses, and other family members. As will be explained later in more detail, some of these parameters are changed in order to produce the aggregate results with regard to wages, profits, and employment shares given by the CGE. The CGE also gives the price vector, which in a last step is used to deflate family income.

Remarks on the Labor Market Specification

The income generation model requires some comments on the assumptions behind its formulation. First of all, despite the availability of data on working time we decided to model the occupational choice as a discrete choice.¹⁵ Secondly, our model assumes that the Vietnamese labor market is segmented along different lines. One line of segmentation separates wage-employment from self-employment. In a perfectly competitive labor market, the returns to labor would be equal for these two types of employment. Yet, segmentation may be justified because income from self-employment is likely to contain a rent from non-labor assets used, and its clearing mechanism may differ from that of wage employment. Information on non-labor assets, land in rural areas and at least a small amount of capital in urban areas, is not available for Vietnam, hence distinct equations need to be estimated even if the labor markets were competitive. In addition, even in those cases where information on non-labor assets is available, a segmented labor market can be justified by the fact that wage-employment may be rationed and self-employment thus “absorbs” those who do not get a job in the preferred wage work. Wage work could be preferred for generating a more steady income stream or for fringe benefits related to this type of employment. Conversely, self-employment might exhibit important externalities, for example for families in which children have to be taken care of. Self-employment of the household head may also create employment opportunities for other family members.

Additional segmentation is assumed within the wage labor market. The segmentation hypothesis along the lines of different gender, skill, and area is strongly supported by the regression results. The same holds for the estimation of different profit functions for agricultural and non-agricultural activities in rural areas.

¹⁵ However, estimating wage equations based on hourly wages did not make a major difference in the coefficients.

Estimation of the occupational choice and earnings equations

As mentioned above, the occupational choice model and the wage and profits equations are estimated in a first step in order to obtain an initial set of coefficients (a_G , β_G , b_F , δ_F , c_H^w , α_H^w , c_H^s , α_H^s) and unobserved characteristics (e_{mi} , ε_{mi} , u_{mi}^w , u_{mi}^s). Unobserved characteristics say for the wage equation can of course only be obtained for those who are actually wage-employed. For self-employed or inactive individuals the unobserved characteristics in the wage-equation are generated by drawing random numbers from a normal distribution. In the same way, we generate unobserved characteristics for the profit function for households in which nobody is self-employed. As we estimate wage and profit functions using ordinary least squares, we assume these unobserved characteristics to be normally distributed. Additionally, unobserved characteristics need to be generated for the occupational choice model. These residuals are assumed to be distributed according to the double exponential law since we estimate a multinomial logit model. They were drawn randomly consistent with the observed occupational choice, i.e. the utility a wage earner relates to wage-employment has to be higher than the utility associated with inactivity or self-employment.

Macro-Micro Links in Detail

As already mentioned, the micro-simulation and the CGE models are linked in a sequential fashion. In a first stage a shock is simulated in the CGE model and then the micro-simulation adjusts micro data so that values for its aggregate variables are consistent with the CGE macro equilibrium. Consistency requires that across the two models the following items are equal: (1) changes in average wages in each segment, (2) changes in average profits in each activity, (3) changes in agricultural output, (4) changes in employment shares in each segment, i.e. the shares of wage-earners, self-employed, and inactive individuals per segment, and (5) the food and non food commodities price changes. The CGE is initially calibrated in such a way that it is consistent with the benchmark micro-simulation. This benchmark micro-simulation is produced by using the set of initial coefficients and unobserved characteristics obtained through the estimation work just described.¹⁶ Formally, the following constraints describe the consistency requirements.

$$\sum_m \sum_{i,g(mi)=G}^{\wedge} IW_{mi} = \quad (7)$$

$$\sum_m \sum_{i,g(mi)=G} Ind[\hat{c}_{h(mi)}^w + z_{mi}\hat{\alpha}_{h(mi)}^w + \hat{u}_{mi}^w > Sup(0, \hat{c}_{h(mi)}^s + z_{mi}\hat{\alpha}_{h(mi)}^s + \hat{u}_{mi}^s)] = E_G$$

$$\sum_m \sum_{i,g(mi)=G} Ind[\hat{c}_{h(mi)}^s + z_{mi}\hat{\alpha}_{h(mi)}^s + \hat{u}_{mi}^s > Sup(0, \hat{c}_{h(mi)}^w + z_{mi}\hat{\alpha}_{h(mi)}^w + \hat{u}_{mi}^w)] = S_G \quad (8)$$

$$\sum_m \sum_{i,g(mi)=G} exp(\hat{a}_G + x_{mi}\hat{\beta}_G + \hat{e}_{mi})IW_{mi} = w_G \quad (9)$$

$$\sum_{m,f(m)=F} exp(\hat{b}_G + z_m\hat{\delta}_G + \hat{\varepsilon}_m)Ind(N_m > 0) = \pi_F \quad (10)$$

¹⁶ By doing this, we simply reproduce the original dataset.

Equation (7) states that, for each labor market segment, the number of wage-employed individuals has to be equal in the CGE (E_G) and micro-simulation systems. “G” stands for the eight labor market segments, i.e. urban male skilled and unskilled, urban female skilled and unskilled, rural male skilled and unskilled, rural female skilled and unskilled labor. The same holds for the number of self-employed in each segment, which is specified in equation (8).

Total wages paid in segment G in the CGE, w_G , have to be equal to the sum of wages over families and wage-employed individuals in the micro-simulation, as indicated by equation (9). This has to be fulfilled also for the profits in activity F as in equation (10). Thus, π_F denotes the total profits for self-employment activity F given by the CGE. The different self-employment activities include urban self-employment, rural agricultural, rural non-agricultural, and rural mixed activities. Note that $\wedge$ indicates that the coefficients, residuals, and indicator function values result from the estimation described above.

A globalization shock produces changes in E_G , the number of wage-employed, S_G , the number of self-employed, w_G , the sum of wages paid in segment G, π_F , the sum of profits paid in activity F, output levels in agriculture x_F , and q , the price vector. The result is a new vector of these variables, which will be identified by an asterisk (E_G^* , S_G^* , w_G^* , π_F^* , x_F^* , q^*). For the above constraints to hold, an appropriate vector of coefficients and prices (a_G , β_G , b_F , δ_F , c_H^w , α_H^w , c_H^s , α_H^s , p) is needed. For the price vector this is trivial, as p equals q . For the other coefficients, many solutions exist and additional constraints have to be introduced. As in Robilliard et al. (2001) our choice is to vary the constants (a_G , b_F , c_H^w , c_H^s) and leave the other coefficients unchanged. We hence assume that the changes in occupational choices and earnings are dependent on personal and household characteristics only to a limited degree. Changing the intercept in one of the wage equations implies that all individuals of the respective segment experience the same increase in log earnings. This increase does not depend on individual characteristics. The same holds for the profit functions. With regard to the occupational choice, it should be noted that the CGE does not allow for distinguishing between the choices of heads, spouses, and others. The changes are thus the same across these groups.

Consistency of the micro-simulation and the CGE requires the solution of the following system of equations. The right hand side variables are those through which the macro model communicates with the micro-simulation. Additionally, the prices for food and non-food items are given by the CGE. However, the price vector is only finally applied in order to deflate household income.

$$\sum_m \sum_{i,g(mi)=G}^{\wedge} IW_{mi} = \sum_m \sum_{i,g(mi)=G} Ind \left[c_{h(mi)}^{*w} + z_{mi} \hat{\alpha}_{h(mi)}^w + \hat{u}_{mi}^w > Sup \left(0, c_{h(mi)}^{*s} + z_{mi} \hat{\alpha}_{h(mi)}^s + \hat{u}_{mi}^s \right) \right] = E_G^* \quad (11)$$

$$\sum_m \sum_{i,g(mi)=G} Ind \left[c_{h(mi)}^{*s} + z_{mi} \hat{\alpha}_{h(mi)}^s + \hat{u}_{mi}^s > Sup \left(0, c_{h(mi)}^{*w} + z_{mi} \hat{\alpha}_{h(mi)}^w + \hat{u}_{mi}^w \right) \right] = S_G^* \quad (12)$$

$$\sum_m \sum_{i, g(mi)=G} \exp(a_G^* + x_{mi} \hat{\beta}_G + \hat{e}_{mi}) \hat{IW}_{mi} = w_G^* \quad (13)$$

$$\sum_{m, f(m)=F} \exp(b_G^* + z_m \hat{\partial}_G + \hat{\varepsilon}_m) \text{Ind}(N_m > 0) = \pi_F^* \quad (14)$$

Equations (11) and (12) require the number of self-employed and wage-employed (and both self-employed and wage-employed in rural areas) to be consistent with the CGE results for each of the eight segments (G). This also holds for the wage equation for each of the segments and the profit function for each of the four activities, as indicated by equations (13) and (14). Hence, the above system contains 28 restrictions. The system has eight unknown constants in the wage equations, four in the profit functions, and 16 in the occupational choice model.¹⁷ Thus we have 28 unknown constants and 28 equations. We obtain the solution by applying standard Gauss-Newton techniques.

Solving the above system gives us a new set of constants (a_G^* , b_F^* , c^{*w}_H , c^{*s}_H), which is then used to compute occupational choices, wages, and profits. The resulting household incomes are deflated by the household group specific price index derived from the CGE results for food and non-food prices.

Linking the CGE and the micro-simulation in the way described above goes beyond simply rescaling various household income sources or reweighing households dependent on the occupation of its members, which is what the RHG approach does. The simulation model takes the different sources of household income into account and mimics individual occupational choices, based on a wide range of individual characteristics, and it is therefore a more accurate method than just rescaling household groups incomes.

An Artificial Panel data set?

At first sight, one may be inclined to think that the simulation method generates a kind of artificial panel, which would be most helpful and interesting from an analytical point of view. If we want to analyze poverty dynamics, we need to trace individuals and households across time. However, to produce a synthetic panel further assumptions need to be introduced. For brevity, the arising problems are illustrated for the case of the wage equation, but they apply to all the simulated relationships. In a dynamic context, the wage equation contains three components. Wages in period 0 consists of observed permanent earnings, i.e. the share of the earnings that can be explained by our model, unobserved permanent earnings e^p and unobserved transitory earnings e^t_0 .

$$\log w_0 = a + x\beta + e = a + x\beta + e^p + e^t_0 \quad (15)$$

From period 0 to period 1, the constant a is modified due to the policy change that triggered the changes in the CGE, so that in the next period we have a^* . If we assume that

¹⁷ Note that the constants of the occupational choice model – though estimated separately for heads, spouses, and others – are changed separately across the eight labour market segments. Therefore, we have 16 unknown constants in the occupational choice model, two occupational choices in each of the four urban labour market segments, and three in each of the four rural segments.

the distribution of the transitory component is the same in both periods, we know that among the people with characteristics x and an unobserved permanent component, e^p , there will be *somebody* with a transitory component equal to e^t_0 . This implies that to any individual in period 0 with earnings given by (15) we may associate *somebody* with earnings given by the following equation.

$$\log w_1 = a^* + x\beta + e = a^* + x\beta + e^p + e^t_0 \quad (16)$$

The individual with earnings given by (16) is not the same as the individual whose earnings were represented by (15). Since this is what we do in the micro-simulation, as set up to this point, we do not generate a synthetic panel, but two cross-sections. Based on two-cross-sections it is of course not possible to trace individuals through time. Yet, there is no problem if we compute aggregate inequality and poverty indicators, which we compare across time. In order to study poverty dynamics though we would have to make sure that we could identify the individuals of the households who cross the poverty line. It is therefore not sufficient to associate *somebody* with unobserved earnings, but a specific individual.

The reason why we cannot simulate a panel arises from the fact that we cannot differentiate between the two unobserved components. However, introducing a set of assumptions with regard to these two terms helps. First, we assume the transitory component to be independent and identically distributed across time. Second, we have to make an assumption about the proportions of the variance of the entire residual term e that is due to the respective components. There are though a number of difficulties related to this method, in particular to the specification of the variance proportions. Some empirical estimates of these proportions can be found in Atkinson et al. (1992) where a number of empirical studies on earnings mobility are surveyed. They find the proportions of the three components in an earnings panel model to differ substantially across different studies. Of course, the total unobserved component is smaller the better the model explains log earnings. The proportion of the transitory component in log earnings covariance varies between less than 10 and 30 percent over long time horizons of more than 10 years. We are not aware of empirical work on earnings mobility in developing countries, which would analyze these issues in detail. There is scope for further research on earnings mobility as some panel datasets have become available. Assuming a small proportion of transitory earnings in developing economies in general may be justified by a number of arguments. Social mobility is generally lower in developing countries.¹⁸ From this, we may infer that transitory earnings account for a smaller proportion of earnings. Additionally, recent research has shown that income shocks remain after a considerable period of time, which also would imply less importance of a transitory component, at least in the short run.¹⁹ On the other hand, the transitory component may be particularly important for small farms, which are exposed to a number of transitory, primarily environmental, risks.

¹⁸ For social mobility in Latin America see Andersen (2000).

¹⁹ See Newhouse (2001) who studies the persistence of transient income shocks to farm households in rural Indonesia. He finds, for example that “about 40 percent of household income shocks remain after four years.”

For the purpose of the poverty transition analysis, we simulated a panel based on the aforementioned assumptions. These panel-based results are of a preliminary character and should be treated with caution, as further research in this field is needed. Experimenting with different proportions in the micro-simulation had a substantial impact on the results. Reducing the proportion of the variance of the residual term e , which is due to the transitory component, to 10 percent produced results in the historical simulation, which were close to those of the original simulation of two cross-sections. Using higher proportions due to the transitory component resulted in considerable increases in inequality indicators. The poverty transition analysis is thus based on the assumption that only 10 percent of the unobserved effects are transitory.²⁰

Household Production Estimation

One of the most common complaints about applying CGE modelling to developing countries is that agricultural production bears little resemblance to a neoclassical/leontief specification of production. In this framework, a single representative production activity (for each agricultural product or sector) combines constant share inputs with factors of production recruited from national labor, capital and (sometimes) land markets. This activity then maximizes profits and some portion of its factor income accrues to rural households who make their expenditure decisions independently.

In reality, most rural sector households in the developing world (and over 50% in Vietnam) are combined production and consumption units. Their factors largely belong to them and a significant part of their output is retained for own consumption. Although evidence indicates that even remote households are aware of local market signals, more remote areas exhibit higher levels of self-sufficiency. This may be partly because of price biases such as those described above, or because of other behavioural characteristics. The consequence in any case is that more careful specification of rural household behavior is needed to capture this behavior.

While we make no pretence to definitive treatment of this, we believe that joint specification of production and consumption can shed important light on rural economic participation, particularly household decisions in response to changing prices and emergent market opportunities. As was already mentioned, economywide CGE models use (combined leontief and CES) industrial technology at the sectoral level, which bears no resemblance to the microeconomics of household production decisions. Standard micro-simulation techniques, including all the antecedents to our approach, have relied on human capital based specification of household profits. While these components are important to wealth determination, household agricultural technologies are in fact very

²⁰ As mentioned before, aggregate inequality indicators increase under the synthetic panel approach. This increase was more pronounced the higher the share of the transitory component. We understood these results when we had a look at the distribution of unobserved earnings. For lower incomes, the distribution of unobserved earnings is skewed to the right, hence implying relatively high unobserved earnings. For higher incomes, the distribution of the entire residual resembles a normal distribution. If we substitute these unobserved earnings or a portion of it by generated normally distributed unobserved earnings, we thus “redistribute” income from the poor to the rich, thereby increasing inequality.

dependent on the quantity and quality of non-human factors and inputs, and this needs to be estimated more directly.

Households are assumed to be characterized by the representative attribute vectors x_i drawn from the LSMS. This means the sample is stratified into household categories representing a discrete partition of each component in x_i as defined in the previous section. With the characteristics come econometrically estimated functions for significant behavioural relationships, including consumption, household production, and labor supply. Consumption is calibrated to a standard ELES form, whereas we model household production with two alternative specifications. and production arising from a household profit function. For calibrating consumption, we assume households divide output between own use and the market. To calibrate production, we are fortunate to have estimates of both types of household production in the SAM. Activities with no own production in the base data will be assumed to be produced by standard neoclassical enterprises.

Where data are available on own production for market, we specify, estimate, and implement household productions at the micro level and aggregate them consistently with the macro level. At the micro level, the production function $F()$ of equation 2 above is estimated in two ways for comparison. In the first instance, we assume corresponds to a standard multi-input CET-CES-GL profit function (see, e.g. Behrman et al: 1989)

$$\pi(q, H) = \left(\sum_{i=1}^{m+n} \alpha_{ii} q_i^{\epsilon} \right)^{1/\epsilon} + \sum_{i=1}^{m+n} \sum_{j \neq i}^{m+n} \alpha_{ij} q_i^{1/2} q_j^{1/2} + \sum_{i=1}^{m+n} \alpha_{iH} q_i H^{1/2}, \quad \alpha_{ij} = \alpha_{ji}$$

This specification can be estimated by standard methods. A more experimental approach to estimation, one that can take account of indivisibilities in inputs and a variety of unobservables, has been suggested by Aguirregabiria (2003) and uses a multinomial logit approach with

$$\Pi^*(S) = (1/\mu_2) \log \left(\sum_{n=0}^M \left[\sum_{m=0}^M \exp \{ \mu_1 \pi(m, n, S) \} \right]^{\mu_2/\mu_1} \right)$$

where S is the vector of output an input/factor prices.

In the case of labor supply, the traditional micro-macro approach specifies discrete occupational choice calibrated to econometric qualitative choice models. This framework might be intellectually more satisfactory for individual labor supply, but it leads to many analytical problems with model closure and solution in either the two-part or synthetic model approach. By contrast, we beleive our representative household sub-sample justifies use of a continuous specification, a CET function $lh_i(x_i, y)$ that yields a three-dimensional vector of household employment shares, divided between own employment, outside employment, and no employment. Such a functional form would be

easily calibrated to VLSS shares, aggregate employment, and is smoothly responsive to relative wages.

Relevant References

- Alatas, V. and F. Bourguignon. 2000. "The evolution of the distribution of income during Indonesian fast growth: 1980-1996". Mimeo. Princeton University.
- Alderman, H., M. Babita, G. Demombynes, N. Makhatha, and B. " Ozler (2002) 'How small can you go? combining census and survey data for mapping poverty in South Africa.' Journal of African Economies 11, 169–200
- Atkinson, A. B., and A. Brandolini (2001): Promise and Pitfalls in the Use of Secondary" Data-Sets: Income Inequality in OECD Countries," Journal of Economic Literature, 39, 3.
- Banerjee, A., and A. Newman (1993) "Occupational Choice and the Process of Development," Journal of Political Economy, 101, 1, 274-98.
- Banerjee, A., and E. Duflo (2000): Inequality and Growth: What Can the Data Say?," NBER Working paper no. 7793.
- Barro, R., and X. Sala-i-Martin (1991) "Convergence Across States and Regions," Brookings Papers on Economic Activity, no. 1, 107-82.
- Bautista, R. (2000), Agriculture-Based Development: A SAM Perspective on Central Viet Nam, TMD Discussion Paper No. 51, International Food Policy Research Institute (IFPRI), Washington, D.C.
- Booth, A. 1998. "The Impact of the Crisis on Poverty and Equity". ASEAN Economic Bulletin. Vol. 15, No. 3.
- Bourguignon F., M. Fournier, and M. Gurgand (2001), "Fast Development with a Stable Income Distribution:Taiwan," 1979-1994, Review of Income and Wealth (June).
- Bourguignon, F., 1979, "Decomposable Income Inequality Measures", Econometrica, 47: 901-20.
- Bourguignon, F., J. de Melo and Akiko Suwa, 1991, "Modeling the effects of adjustment programs on income distribution", World Development, 19(11): 1527-1544.
- Chesher, A., and C. Schluter (2002): Welfare Measurement and Measurement Error," Review of Economic Studies, forthcoming.
- Cogneau, D. and A.S. Robilliard. 2000. "Growth, Distribution and Poverty in Madagascar: Learning from a Microsimulation Model in a General Equilibrium Framework". IFPRI TMD DP no. 61 and DIAL DT/2001/14.

- Cowell, F., and M.-P. Victoria-Feser (1996) "Robustness Properties of Inequality Measures," *Econometrica*, 64, 1, 77-101.
- Datt, G., and M. Ravallion [1992], "Growth and Redistribution Components of Changes in Poverty Measures: A Decomposition with Applications to Brazil and India in the 1980s." *Journal of Development Economics* v38, n2 (April 1992): 275-95
- Datt, G., and T. Walker [2002], *PovStat 2.12, A Poverty Projection Toolkit, User's Manual*, The World Bank, October 14, 2002, mimeo.
- Datt, G., Krishnan Ramadas, Dominique van der Mensbrugghe, Thomas Walker, and Quentin Wodon, [2003] "Predicting the effect of aggregate growth on poverty," The World Bank, Mimeo.
- Deaton, A. (1997): *The Analysis of Household Surveys: A Microeconometric Approach to Development Policy*. Washington, D.C.: The Johns Hopkins University Press for the World Bank.
- Deininger, K., and L. Squire (1996): A New Data Set Measuring Income Inequality," *The World Bank Economic Review*, 10, 565-91.
- Demombynes, G., C. Elbers, J. O. Lanjouw, P. Lanjouw, J. Mistiaen, and, B. Ozier (2002): *Producing an Improved Geographic Profile of Poverty: Methodology and Evidence from Three Developing Countries*," WIDER Discussion Paper no. 2002/39, The United Nations.
- Elbers, C., and P. Lanjouw (2001), "Intersectoral Transfer, Growth, and Inequality in Rural Ecuador," *World Development*, 29, 3, 481-96.
- Elbers, C., J.O. Lanjouw, and P. Lanjouw (2003) 'Micro-level estimation of poverty and inequality.' *Econometrica* 71(1), 355–364
- Elbers, C., J. O. Lanjouw, and P. Lanjouw (2000): *Welfare in Villages and Towns: Micro-Measurement of Poverty and Inequality*," Tinbergen Institute Working Paper no. 2000-029/2.
- Elbers, C., J. O. Lanjouw, P. Lanjouw, and P. G. Leite (2002), "Poverty and Inequality in Brazil: New Estimates from Combined PPV-PNAD Data," Unpublished Manuscript, The World Bank.
- Elbers, C., P. Lanjouw, J. Mistiaen, B. " Ozler, and K. Simler (2002) 'Are neighbours equal? estimating local inequality in three developing countries.' Paper presented

- at the Cornell/LSE/WIDER Conference on Spatial Inequality and Development, The World Bank. Downloadable from <http://www.unu.wider.edu>
- Fields, G. (2001), "Economic Growth and Inequality: A Review of the Empirical Evidence," Chapter 3 in *Distribution and Development: A New Look at the Developing World*. Russel Sage Foundation and MIT Press.
- Foster, J., J. Greer and E. Thorbecke, 1984, "A Class of Decomposable Poverty Measures", *Econometrica* 52: 761-5.
- Frankenberg E., D. Thomas, and K. Beegle. 1999. "The Real Cost of Indonesia's Economic Crisis: Preliminary Findings from the Indonesia Family Life Surveys. International Labour Organization (ILO). 1998. *Employment Challenges for the Indonesian Economic Crisis*. Jakarta: ILO and United Nations Development Program.
- Ghosh, M., and J. N. K. Rao (1994): "Small Area Estimation: An Appraisal," *Statistical Science*, 9, 55-93.
- GSO (2000), *Vietnam Living Standards Survey*, General Statistics Office, Hanoi.
- GSO (2002a), *State Budget: Final Accounts for 2000*, General Statistics Office, Hanoi.
- GSO (2002b), *The Results of Enterprise Census of 1 April 2001*, Hanoi Statistical Publishing House, Hanoi.
- GSO (2003a), *Input-Output Tables for Vietnam*, General Statistics Office, Hanoi.
- GSO (2003b), *The System of National Accounts for Vietnam*, General Statistics Office, Hanoi.
- GSO (2003c), *Data for 2000 Provided to Supplement The System of National Accounts for Vietnam*, Mimeo, General Statistics Office, Hanoi.
- Hall, P. 1992. *The Bootstrap and Edgeworth Expansion*. New York: Springer-Verlag.
- Hall, P. and J.L. Horowitz 1995. "Bootstrap critical values for tests based on generalized-method-of moments estimators," *Econometrica*.
- Hellerstein, J., and G. Imbens (1999): "Imposing Moment Restrictions from Auxiliary Data by Weighting," *Review of Economics and Statistics*, 81, 1, 1-14.
- Huong, L. (2000), *Impacts of Trade and Investment Policy on Income Distribution in Vietnam Using a General Equilibrium Framework*, ANU thesis, Canberra.

- IFPRI (2003), *Income Diversification and Poverty in the Northern Uplands of Vietnam*, Report prepared for the Social Development Division, Japan Bank for International Development, Washington and Tokyo.
- Kanbur, R., 1984, "The Measurement and Decomposition of Inequality and Poverty" in: *Mathematical Methods in Economics*, Chapter 16, F van der Ploeg (ed.), John Wiley and Sons.
- Keyzer, M. (2000): Reweighting Survey Observations by Monte Carlo Integration on a Census," Stichting Onderzoek Wereldvoedselvoorziening, Statistical Working Paper no. 00.04, the Vrije Universiteit, Amsterdam.
- Levinsohn, J., S. Berry, and J. Friedman. 1999. "Impacts of the Indonesian Crisis: Price Changes and the Poor". NBER Working Paper No. 7194. Cambridge, MA: NBER.
- Manning, C. 2000. "Labour Market Adjustment to Indonesia's Economic Crisis: Context, Trends and Implications". *Bulletin of Indonesian Economic Studies* 36:1.
- Minot, N. (2000) 'Generating disaggregated poverty maps: An application to Vietnam.' *World Development* 28(2), 319–331
- Nielsen, C. P. (2001), *Social Accounting Matrices for Vietnam 1996 and 1997*, Working Paper No. 8, Danish State Institute of Agricultural and Fisheries Economics (SJFI), Copenhagen.
- Pakes, A., and D. Pollard (1989): Simulation and the Asymptotics of Optimization Estimators," *Econometrica*, 57, 1027-58.
- Plumb, M. (2001), *Empirical tax modeling: an applied general equilibrium model for the UK incorporating micro-unit household data and imperfect competition*, Dphil Thesis, Nuffield College, University of Oxford.
- Pradhan, M., A. Suryahadi, S. Sumarto, and L. Pritchett. 2000. "Measurements of Poverty in Indonesia: 1996, 1999, and Beyond". Research Working Paper. SMERU.
- Rao, J. N. K. (1999): Some Recent Advances in Model-Based Small Area Estimation," *Survey Methodology*, 25, 175-86.
- Ravallion, M. (1998): Does Aggregation Hide the Harmful Effects of Inequality on Growth?," *Economics Letters*, 61, 1, 73-7.

- Ravallion, M., 1996, "Issues in measuring and modeling poverty", World Bank Policy Research Working Paper 1615
- Ravallion, M., and K. Chao (1989) 'Targeting policies for poverty alleviation under imperfect information: Algorithms and applications.' *Journal of Policy Modeling* 11(2), 213–224
- Ravallion, M., and S. Chen, 1997, "What Can New Survey Data Tell Us about Recent Changes in Distribution and Poverty?", *World Bank Economic Review*, 11(2): 357-82.
- Robilliard, Anne-Sophie, and François Bourguignon, Sherman Robinson, "Examining the Social Impact of the Indonesian Financial Crisis using a Micro -Macro Model," paper presented at the European ABCDE Conference of the World Bank, Bergen, Norway.
- Tarozzi, A. (2002), "Estimating Comparable Poverty Counts from Incomparable Surveys: Measuring Poverty in India," RPDS Working paper no. 213, Princeton University.
- Tarp, Finn, David Roland-Holst, John Rand, Pham Lan Huong, Vo Tri Thanh, Pham Hoang Ha and Do Thi Thu Huong. (2001) *A 1999 Social Accounting Matrix for Vietnam*. The Gioi Publishers. Hanoi.
- Traub, J.F., and A.G. Werschulz (1998), *Complexity and Information*. Cambridge: Cambridge University Press.
- United Nations (undated), *Social Accounting Matrix Based Macromodels for Policy Analysis in Developing Countries*, United Nations, New York.