BACKGROUND PAPER

FOR THE WORLD DEVELOPMENT REPORT 2008

Estimating the Effects of Aggregate Agricultural Growth on the Distribution of Expenditures

Ethan Ligon and Elisabeth Sadoulet

ESTIMATING THE EFFECTS OF AGGREGATE AGRICULTURAL GROWTH ON THE DISTRIBUTION OF EXPENDITURES

ETHAN LIGON AND ELISABETH SADOULET

Over the last several decades, the World Bank has accumulated a large number of datasets from a large number of countries which are based on household-level surveys, statistically representative of the populations of those countries, and which include data on non-durable expenditures. These data on expenditures can be used to measure economic welfare—indeed, this kind of measurement is a chief *raison d'etre* of this collection of survey data.

Though the micro-data from these surveys are not generally available, the Bank provides data on aggregate expenditures by decile for many of these countries. Further, for many countries data from more than one year is available, so that it's possible to construct an unbalanced panel of data on the level and distribution of expenditures for a number of countries over the last several decades. We also have data on country-level measures of agricultural income, as well as other aggregate income. The question: how do changes in the sectoral composition of income affect the distribution of expenditures across households within a country?

1. Models

The question of how changes in the sectoral composition of income affect the distribution of expenditures is an important one for all many of policy issues. Given this importance, it's surprising how little reliable guidance there seems to be in either the theoretical or empirical literature. Here we briefly and selectively review a few models and bits of evidence the subject. We will assume two sectors throughout—an agricultural and non-agricultural sector, both for simplicity, and because

Date: All code and data used to generate this paper are available at `http://code.google.com/p/inequalitygrowth/source`. Last committed as Revision: 667 on Date: 2007-04-10 00:35:50 -0700 (Tue, 10 Apr 2007) .

Many thanks to Alain de Janvry for suggesting this topic, and to numerous helpful researchers at the World Bank for providing the data. Fangwen Lu provided invaluable research assistance.

this assumption is consistent with our own empirical work reported below.

The benchmark models employed in the trade literature often focus on the impacts of sectoral changes on incomes, but the transmission of changes in sectoral composition to changes in the distribution of expenditures is seldom contemplated (Davis and Mishra, 2007). To the extent that the matter is considered, the usual assumption in the trade literature is that expenditures will be equal to income. Thus, to trace out the impact of changes in the sectoral composition of income on the distribution of expenditures, one could start with data on the distribution of employment conditional on position in the expenditure distribution. An increase in income in one sector would affect the welfare only of the part of the population actually employed in that sector. If the expenditures are distributed differently across households in the two sectors, then an increase in income in one sector will have an effect on the aggregate distribution. To take a particularly germane example: Suppose that households employed in the agricultural sector tend to be poorer, so that they are disproportionately featured in the left-hand tail of the exenditure distribution. Then if an increase in agricultural income increases the expenditures of agricultural households, then it will also have an equalizing effect on the entire distribution of expenditures (Thorbecke and Jung, 1996).

The effect of an increase in agricultural income on distribution need not match this prediction, however. If, for example, workers can costlessly change sectors, then we'd predict that an increase in agricultural income would instead accrue to the owners of immobile factors involved in agricultural production (e.g., land). Any shock which increased the marginal product of labor in agriculture relative to the its marginal product in other pursuits would stimulate an increase in the share of agricultural employment, rather than in relative agricultural wages.

However, even if workers are mobile and wages are equated across sectors, differences in the rate of growth of different sectors can result in changes in the distribution of expenditures. For example, Loayza and Raddatz (2006) formulate a model in which expenditures of the poor are equal to the prevailing wage, while non-poor households can borrow or lend to smooth away the effects of variation in labor income on expenditures (alternatively, one could assume that the non-poor are the owners of the economy's capital stock). The question of how sectoral growth effects poverty then boils down to its effects on real wages. Not surprisingly, the model shows that these effects are larger for sectors with larger employment and a lower elasticity of demand for labor. Using a cross-sectional dataset of country-level aggregates, Loayza and Raddatz find evidence that growth in the income of sectors with high labor shares has a disproportionate effect in reducing poverty rates.

2. Methods

Index the set of countries in our dataset by $\ell = 1, 2, \ldots, L$, and index time by $t = 1, 2, \ldots, T$. Let q index expenditures quantiles (deciles in this application), and let (ℓ, q) denote the qth expenditures quantile in country ℓ.

The value of expenditures for quantile q in country ℓ at time t is denoted by $c_t^{(\ell,q)}$. Aggregate agricultural income in the same country and same year is denoted by $y_{\ell t}^1$; non-agricultural income by $y_{\ell t}^2$.

Now, consider the estimating equation

$$\Delta \log c_t^{(\ell,q)} = \alpha^{(\ell,q)} + \eta_t + \eta_t^\ell + \beta_q^1 \Delta \log y_{\ell t}^1 + \beta_q^2 \Delta \log y_{\ell t}^2 + \epsilon_t^{(\ell,q)}. \tag{1}$$

Here the terms $\{\alpha^{(\ell,q)}\}$ are country-quantile "fixed effects" which capture variation in differences in the expected growth 'trend' of log expenditures across countries and deciles, but not across time. The terms $\{\eta_t\}$ capture the average impact of common global shocks on all country-decile-years, while the terms $\beta_q^1 \Delta \log y_{\ell t}^1$ and $\beta_q^2 \Delta \log y_{\ell t}^2$ capture, respectively, the effects of a changes in the growth of agricultural and non-agricultural income on the growth of expenditures of decile q within the country.

3. Issues

3.1. **Accounting for an Unbalanced Panel.**

The panel we're working with for this problem is quite unbalanced. Table 1 gives a list of countries and years for which we have usable data. Note that to estimate (1) we have to have at least three years of data—three years to get two differences, and two differences to estimate the country-quantile fixed effects $\{\alpha^{(\ell,q)}\}$.

Country Name	Years of Available Data
Armenia	1999, 2001–2003
Bangladesh	1984, 1986, 1989, 1992, 1996, 2000
Belarus	2000–2002
Bulgaria	1989, 1997, 2001, 2003
Burkina Faso	1994, 1998, 2003
Cote dIvoire	1985–1988, 1993, 1995, 1998, 2002
Croatia	1998–2001

Continued on next page

Table 1 – continued from previous page

Country Name	Years of Available Data
Estonia	1995, 1998, 2001–2003
Georgia	1996–2003
Ghana	1988–1989, 1992, 1998
Hungary	1998–1999, 2001–2002
India (rural)	1978, 1983, 1986–1990, 1992–1997, 1999
India (urban)	1978, 1983, 1986–1990, 1992–1997, 1999
Indonesia	1987, 1993, 1996, 1998–2000, 2002
Iran, Islamic Rep.	1986, 1990, 1994, 1998
Jordan	1987, 1992, 1997, 2003
Kenya	1992, 1994, 1997
Kyrgyz Republic	1993, 1997–2003
Lao PDR	1992, 1997, 2002
Latvia	1998, 2002–2003
Macedonia, FYR	1998, 2000, 2002–2003
Mali	1989, 1994, 2001
Mauritania	1987, 1993, 1996, 2000
Moldova	1997–1999, 2001–2003
Mongolia	1995, 1998, 2002
Morocco	1985, 1991, 1999
Nicaragua	1993, 1998, 2001
Niger	1992, 1994–1995
Nigeria	1986, 1993, 1997, 2003
Pakistan	1987, 1991, 1993, 1997
Poland	1992, 1996, 2000–2002
Romania	1998, 2000, 2002–2003
Russian Federation	1993, 1996, 1998, 2000–2002
Senegal	1991, 1995, 2001
South Africa	1993, 1995, 2000
Thailand	1981, 1988, 1992, 1996, 1998–2000, 2002
Tunisia	1985, 1990, 1995, 2000
Turkey	1987, 1994, 2000, 2002–2003
Uganda	1989, 1992, 1996, 1999, 2002
Ukraine	1995–1996, 2002–2003
Vietnam	1993, 1998, 2002, 2004
Zambia	1991, 1993, 1996, 1998, 2003

Table 1: Countries and years included in the analysis.

3.2. **Dealing with Endogeneity.** Unobserved shocks which influence either the level or the distribution of expenditures within a country may also influence aggregate sources of income. For example, standard accounts of the determination of aggregate expenditures at the country-level would assign a key role to country-level variation in prevailing interest rates, but these same interest rates play a key role in determining investment and hence income from both agricultural and non-agricultural sources. Interest rates, in turn, will depend on the countries monetary or exchange rate management. A more extreme example might be taken from current political turmoil in Zimbabwe, where repressive political measures taken against farmers have at once produced a vast reduction in agricultural income and an extensive system of price controls, affecting expenditures.

To deal with this problem, we adopt a simple instrumental variables strategy to deal with the potential endogeneity of both agricultural and non-agricultural sources of income. The idea is simple; we use the mean of *neighboring* countries' growth rates of agricultural income as an instrument for own agricultural income growth (a neighboring country is defined as one which shares a common border). The idea is that many of the unobserved shocks which might simultaneously influence income and expenditures will be country-specific, while at last some of the shocks which influence agricultural productivity (e.g., weather related shocks) are likely to be correlated across neighboring countries.

4. Experiments

Here we try a variety of specifications and minor modifications to data, estimation, and so on.

4.1. **Benchmark Specification.** In this experiment we adopt an instrumental variables estimator compute point estimates of the effects of agricultural income growth on the distribution of expenditure growth across the population. We assume a restricted version of (2) for our main equation of interest

$$\Delta \log c_t^{(\ell,q)} = \alpha^{(\ell,q)} + \beta^1 \Delta \log y_{\ell t}^1 + \beta^2 \Delta \log y_{\ell t}^2 + \epsilon_t^{(\ell,q)}, \tag{2}$$

where the left-hand side variable is the change over time of the logarithm of expenditures at time t for quantile q in country ℓ, where the right-hand side variable $\alpha^{(\ell,q)}$ denotes a country-quantile fixed effect, and where the terms $\beta^i \Delta \log y_{\ell t}^i$ $(i = 1, 2)$ capture the effects of income growth from both agricultural and non-agricultural sources in country ℓ at time t on expenditure growth and distribution.

To deal with the possibly endogenous income variables, we also make use of information on average income growth in neighboring countries, which we denote $\Delta \log y^i_{-\ell t}$. These neighboring income growth variables are taken to be related to own income growth via

$$\Delta \log y^i_{\ell t} = \gamma^i \Delta \log y^i_{-\ell t} + \eta^i_t + v^{i,\ell}_t. \tag{3}$$

The parameters of this equation are estimated via least squares, so that the residuals $v^{i,\ell}_t$ are orthogonal to the right-hand side variables by construction. In addition to these orthogonality conditions, the estimator exploits the conditions that

$$\mathrm{E}(\epsilon^{(\ell,q)}_t | \Delta \log y^1_{-\ell t}, \Delta \log y^2_{-\ell t}, \mathbf{1}_t, \mathbf{1}_{(\ell,q)}) = 0 \tag{4}$$

where the notation $\mathbf{1}$ denotes an indicator variable which varies according to the variable subscript (so that $\mathbf{1}_t$, for example, describes a collection of year dummies). The estimator also exploits

$$\mathrm{E}(v^{i,\ell}_t | \Delta \log y^1_{-\ell t}, \Delta \log y^2_{-\ell t}, \mathbf{1}_\ell) = 0. \tag{5}$$

In addition, to conduct hypothesis testing and inference we assume that the residuals $\{\epsilon^{(\ell,q)}_t\}$ are homoskedastic and independently distributed (we relax both of these assumptions in Experiment 4.2 below).

One way of implementing this estimator is to do so in two stages. Our first stage regression involves regressing this growth rate on the average growth rate of agricultural income in neighboring countries (as described in Section 3.2), along with a collection of year dummies. We similarly regress the growth rate of *non*-agricultural income on the average growth rate of non-agricultural income in neighboring countries and a collection of year dummies.

Following this procedure gives rise to the results reported in the second column and top panel of Table 2. The coefficient estimates reported in the table have the interpretation of elasticities; thus, our first stage estimates here imply that a growth rate of ten percent in neighboring countries' agricultural income will increase country i's agricultural growth rate by roughly two percent. Though this coefficient isn't significant (at conventional levels of statistical confidence) in this regression, along with the collection of year effects it is jointly significant. Further, its correlation with these year effects is not so large as to change the estimated coefficient very much from a specification in which the year effects are omitted, and replaced with a constant (reported in the first column of the top panel). Still, the relatively low R^2 of 18.5 per cent in this first stage makes us wish for a stronger instrument.

	Border Instrument		
	Constant	Year Effects	Country Effects
Observations	2060	2060	2060
Agriculture			
Coefficient Est.	0.211*	0.198	0.135
Std. Errors	(0.122)	(0.124)	(0.126)
R^2	0.018	0.185	0.166
Non Agriculture			
Coefficient Est.	0.584***	0.458***	0.541***
Std. Errors	(0.090)	(0.097)	(0.100)
R^2	0.204	0.317	0.507

TABLE 2. First stage regression of the growth rate of agricultural income on the average of neighboring countries' growth rates of agricultural income. Different columns reflect different error-correction strategies. The first column includes only a constant; the second a collection of year-dummy variables; and the third a collection of country-fixed effects.

Our other 'first stage' regression, of the growth rate of non-agricultural income on neighbors' growth rates on non-agricultural income, is reported in the bottom panel of Table 2. Here the estimated point elasticities are higher, ranging from 0.458 when we include a collection of year dummies in the regression to 0.584 when we only include a constant. In each of these three specifications the estimated coefficient is highly significant. Further, the fit of these regressions is much better—including only a constant and the growth term gives an R^2 statistic of 20 per cent, compared to 1.8 per cent for the case of agricultural income growth.

The question we're interested in answering, however, is not really how innovations in income growth are transmitted across borders, but rather how these innovations influence growth in expenditures across population deciles. To address this question we use a second stage of estimation, which involves constructing predictions $\widehat{\Delta \log y_{it}}$ of agricultural income growth and non-agricultural income growth, and then replacing $\Delta \log y_{it}$ with these predicted values in our estimating equation.

Some tentative answers emerge in Table 3. The first thing worthy of note in this table is actually just a some simple summary statistics, labeled "Shares" in the table. That is, the average share of agricultural

Deciles	Agricultural Income Growth	Non-Agricultural Income Growth
Shares	0.226	0.774
Std. Errors	0.134	0.248
10%	1.652***	−0.708**
20%	0.833***	0.796**
30%	0.620***	1.000***
40%	0.508***	1.113***
50%	0.439**	1.176***
60%	0.392**	1.237***
70%	0.356**	1.265***
80%	0.305**	1.271***
90%	0.223*	1.229***
100%	−0.233*	1.154***

Table 3. Second stage regression of the growth rate of decile expenditures on the aggregate growth rates of agricultural and non-agricultural income. The annotations *, **, and *** indicate significance of the corresponding coefficient estimate with levels of confidence corresponding to 90%, 95%, and 99%. In addition to the reported variables, the (second stage) estimating equation includes fixed effects for each country-decile, while the first stage includes year effects.

income out of total income across all the country-years in our sample is 22.6 percent. A simple arithmetical consequence of this fact is that, distributional effects aside, the average country would always prefer to see one percent growth in non-agricultural sources of income rather than one per cent in agricultural sources, simply because this non-agricultural income accounts for a much larger share of total income.

However, when taking distributional effects into account the matter becomes much less clear. The reason for this is that growth in agricultural income has a much larger impact on expenditure growth for poorer households than does the growth of non-agricultural income. In fact, our estimates suggest that a one per cent increase in aggregate agricultural income will accrue disproportionately to the poorest decile, as expenditures for these households increase by an estimated 1.65 per cent in response. In contrast, though a one percent increase in the size of non-agricultural income will, on average, have a much larger effect on *total* income and expenditures, it will actually have a predicted *negative* effect on the expenditures of the poorest decile, with these households experiencing on average a 0.7 per cent *decrease* in their expenditures.

Households in the rest of the expenditure distribution show a similar—if less extreme—pattern. As the wealth of the decile increases we see a monotonic decline in the elasticity estimated for agricultural income growth and (with a single small exception for the top two deciles) a monotonic increase in the estimated elasticity for non-agricultural sources of income. The third decile has an estimated elasticity of almost exactly one, while wealthier households benefit more than proportionally from increases in non-agricultural income growth.

Though these results 'make sense,' the particular specification and restrictions we use are certainly not inevitable or obviously correct. Issues of particular importance include the validity of the income growth instruments we employ, and the particular form of what we'll call our "error correction" strategy (language we borrow from Amemiya (1985)). There's ordinarily a tradeoff in panel estimators involving the addition of error correction variables (fixed effects, year effects, and so on) between the consistency of the estimator and a reduction in precision and the power of inference. This trade-off is perhaps particularly acute for these data, as the very small, unbalanced panel makes the use of error correction at once more important and yet more costly.

4.2. **Benchmark Specification with Robust Standard Errors.** Here we start with exactly the specification and two-stage estimator of Experiment 4.1, but make corrections to our estimated standard errors by following a procedure suggested by Arellano (1987) which relaxes the assumptions of independence and homoskedasticity relied on to estimated standard errors of our point estimates above. To be precise, let ϵ_t^ℓ denote the Q-vector of residuals for all the quantiles in country ℓ and time t. Let the index set T_ℓ denote the set of years (excepting the first) for which we have data for country ℓ. Further, let the total number of country-periods be equal to $\bar{T}+1$. Then we estimate the $Q \times Q$ covariance matrix of these residuals by

$$\hat{\Sigma} = \bar{T}^{-1} \sum_{\ell=1}^{L} \sum_{t \in T_\ell} \epsilon_t^\ell (\epsilon_t^\ell)',$$

which converges to $\mathrm{E}\epsilon_t^\ell(\epsilon_t^\ell)'$ under a weak law of large numbers. The idea is that estimating this matrix allows for arbitrary forms of correlation and heteroskedasticity across deciles in a given country-year. It does not allow for serial correlation, but given the short and variable

	Sectoral Income Growth		
Deciles	Agricultural	Non-Agricultural	Test
10%	1.652^{***}	-0.708	33.406
	(0.224)	(0.450)	(0.000)
20%	0.833^{***}	0.796^{**}	0.010
	(0.203)	(0.377)	(0.922)
30%	0.620^{***}	1.000^{**}	1.100
	(0.220)	(0.434)	(0.294)
40%	0.508^{**}	1.113^{***}	3.596
	(0.222)	(0.343)	(0.058)
50%	0.439^{**}	1.176^{***}	5.734
	(0.214)	(0.384)	(0.017)
60%	0.392^{*}	1.237^{***}	5.015
	(0.201)	(0.379)	(0.025)
70%	0.356	1.265^{***}	7.359
	(0.244)	(0.404)	(0.007)
80%	0.305	1.271^{***}	8.304
	(0.206)	(0.377)	(0.004)
90%	0.223	1.229^{***}	9.258
	(0.256)	(0.370)	(0.002)
100%	-0.233	1.154^{***}	42.599
	(0.220)	(0.267)	(0.000)

TABLE 4. Second stage regression of the growth rate of decile expenditures on the aggregate growth rates of agricultural and non-agricultural income. In the first two columns the annotations *, **, and *** indicate significance of the corresponding coefficient estimate with levels of confidence corresponding to 90%, 95%, and 99%. In addition to the reported variables, the (second stage) estimating equation includes fixed effects for each country-decile, while the first stage includes year effects. Reported standard errors are robust to arbitrary forms of heteroskedasticity, and arbitrary patterns of correlation across deciles. The final column reports chi^2 statistics associated with the null hypothesis that the coefficients reported in the first two columns are equal. Parenthetical numbers in the final column are p-values associated with this test of equality.

nature of the time-series available to us any attempts to estimate serial correlation would presumably yield unreliable estimates in any case.

Let $X = [\Delta \log y^1_{\ell t}, \Delta \log y^2_{\ell t}, \imath_\ell] \otimes I_Q$ be the "right-hand-side" variables in the estimating equation (2) described in Experiment 4.1, and similarly let $Z = [\Delta \log y^1_{-\ell t}, \Delta \log y^2_{-\ell t}, \imath_t]$ denote the "right-hand-side" variables in the first stage regression. We construct an estimator of the covariance matrix of the parameter estimates from the second-stage regression

$$\hat{V} = [X'Z(Z'(I_{\bar{T}} \otimes \hat{\Sigma})Z)^{-1}Z'X]^{-1}.$$

It is this estimator that yields the standard errors which appear in Table 4. The magnitude of these estimated standard errors is roughly twice the magnitude of the standard errors estimated under the hypothesis of independence and homoskedasticity; still, a preponderance of the estimated coefficients are significant at conventional levels of significance.

4.3. **Ordinary Least Squares.** The results we feature above in Experiment 4.1 implement methods to deal with the potential endogeneity of both agricultural and non-agricultural income growth. While we think the logic suggesting that both income and expenditures are apt to be simultaneously determined by country-specific shocks is compelling, it remains to be seen whether or not evidence of the endogeneity of income is strong enough to motivate the instrumental variables approach we take to estimation.

Accordingly, in this experiment we pursue the simpler alternative of estimating the parameters of (2) via ordinary least squares. Results are given in Table 5. The flavor of the results bears a similarity to the results reported in Experiment 4.1, in that the poorest decile of households appears to benefit more from agricultural income growth than non-agricultural income growth, but that these benefits from agricultural income growth generally fall with the level of expenditures, while the benefits of non-agricultural income growth start off small for the poorest households, but tend to increase with household expenditures.

Despite this similarity, a Hausman test easily rejects the null hypothesis that the coefficient estimates from the OLS estimator are equal to the estimates from Experiment 4.1 (the χ^2_{440} statistic in this case is 1639.0), providing statistical support for the conjecture that income growth is endogenous.

Deciles	Agriculture Income Growth	Non-Agriculture Income Growth
Shares	0.226	0.774
Std. Errors	0.123	0.227
10%	0.755***	0.432*
20%	0.213*	0.865***
30%	0.162	0.916***
40%	0.141	0.962***
50%	0.134	0.990***
60%	0.129	1.025***
70%	0.128	1.048***
80%	0.128	1.064***
90%	0.133	1.065***
100%	−0.006	1.100***

TABLE 5. OLS regression of the growth rate of decile expenditures on the aggregate growth rates of agricultural and non-agricultural income. The annotations *, **, and *** indicate significance of the corresponding coefficient estimate with levels of confidence corresponding to 90%, 95%, and 99%. In addition to the reported variables, the estimating equation includes fixed effects for each country-decile.

4.4. **Comparing Elasticities for Countries with High and Low Agricultural Income Shares.** This experiment modifies the estimation strategy of Experiment 4.1 by permitting the estimated elasticities associated with agricultural and non-agricultural income growth to differ depending on whether the country in question is has a higher or lower initial share of agricultural income.

The initial income shares are calculated as follows. For twenty-five countries out of a total of forty-two, the initial shares are calculated using the four years of data from 1980 to 1983. For all of the remaining countries, data are not available as early as 1980. We'd still like the earliest available data on income shares for these countries. Accordingly, for sixteen of the remaining countries except Poland, we use the earliest available four year period. The four years are picked in a way such that we can not only calculate the exogenous (to our decile samples) income shares, but also we can compare income shares across counties according to similar bases. Although using the earliest available data is a common strategy for avoiding endogeneity, it will induce data selection problems in our context, because some countries have very early data available while others not, and countries tends to have larger agriculture income shares in earlier years. We calculate the income shares for each year, and then take the average of the four years. For Poland, we only use data from 1992 because the earliest available income data for Poland is in 1992 and our decile sample for Poland also starts in 1992. [1]

An alternative way of calculating shares is to calculate current shares based on information from 1990 to 2003. Comparing the two different calculations, Mongolia, Iran, and Belarus changed from initially low-share countries to high-share countries more recently, while Nicaragua, Indonesia and Mauritania change in the other direction.

The estimation is conducted in two stages, as discussed in Experiment 4.1, and exploits a similar set of restrictions. In the first stage, we predict aggregate agricultural and non-agricultural income at the country-year level using corresponding exogeneous variables, exactly as described in Experiment 4.1. In the second stage, we interact predicted agricultural income and non-agricultural income with both the high/low variable and with decile dummies. This yields the results reported in Table 6.

Results once again bear a broad similarity to the results reported in Experiment 4.1. For countries with either high or low initial income shares, the effects of agricultural income growth on the growth of expenditures of the poorest decile are large, positive, and significant, but fall with household expenditure levels, while the effects of non-agricultural income growth tend to show the opposite pattern. Comparing countries with low initial shares to countries with high initial shares, we see that the rate at which the estimated elasticities associated with agricultural income growth fall is lower for countries with higher initial shares. However, the deciles across these two groups of countries are not really strictly comparable—bear in mind that households in the bottom decile of the "high" group of country are considerably more poor than their counterparts in the "low" group.

4.5. **Instrumental Variables Regression.** Here we consider the consequences of relaxing some of the restrictions we employ in the benchmark estimator discussed in Experiment 4.1. We start with exactly the

[1]Available income data mean, for a country-year, all of the following five variables are available: Agriculture, value added (constant LCU), Agriculture, value added (current LCU), GDP (constant LCU), GDP (current LCU) and Purchasing power parity conversion factor (LCU per international $)).

Deciles	Agriculture Income Growth	Non-Agriculture Income Growth
Shares	0.226	0.774
Std. Err.	0.166	0.372
L10%	2.583***	−2.081***
L20%	0.479**	0.443
L30%	0.162	0.674*
L40%	0.019	0.807**
L50%	−0.069	0.851**
L60%	−0.134	0.897**
L70%	−0.191	0.904**
L80%	−0.259	0.898**
L90%	−0.367**	0.860**
L100%	−1.082***	0.912**
H10%	1.181***	0.336
H20%	1.040***	1.062**
H30%	0.884***	1.244***
H40%	0.787**	1.343***
H50%	0.729**	1.420***
H60%	0.694**	1.491***
H70%	0.670**	1.536***
H80%	0.628**	1.551***
H90%	0.562**	1.505***
H100%	0.244	1.334***

TABLE 6. Second stage regression of the growth rate of decile expenditures on the aggregate growth rates of agricultural and non-agricultural income. Expenditure and income variables are interacted with variables that indicate whether or not the country had high or low initial share of agricultural income in total income.The annotations *, **, and *** indicate significance of the corresponding coefficient estimate with levels of confidence corresponding to 90%, 95%, and 99%. In addition to the reported variables, the (second stage) estimating equation includes fixed effects for each country-decile, while the first stage includes year effects.

same basic estimating equation,

$$\Delta \log c_t^{(\ell,q)} = \alpha^{(\ell,q)} + \beta^1 \Delta \log y_{\ell t}^1 + \beta^2 \Delta \log y_{\ell t}^2 + \epsilon_t^{(\ell,q)}$$

Deciles	Agriculture Income Growth	Non-Agriculture Income Growth
Shares	0.226	0.774
Std. Errors	1.531	0.997
10%	-2.854^{*}	-2.669^{**}
20%	0.337	0.756
30%	0.709	1.170
40%	0.834	1.367
50%	0.896	1.477
60%	0.897	1.549
70%	0.898	1.589
80%	0.854	1.584
90%	0.784	1.515
100%	1.769	1.481

TABLE 7. Instrumental variables regression of the growth rate of decile expenditures on the aggregate growth rates of agricultural and non-agricultural income. The annotations *, **, and *** indicate significance of the corresponding coefficient estimate with levels of confidence corresponding to 90%, 95%, and 99%. In addition to the reported variables, the estimating equation includes fixed effects for each country-decile year effects.

and exploit the same moment restrictions

$$\mathrm{E}(\epsilon_t^{(\ell,q)}|\Delta \log y^1_{-\ell t}, \Delta \log y^2_{-\ell t}, \mathbf{1}_t, \mathbf{1}_{(\ell,q)}) = 0.$$

However, in Experiment 4.1 we also exploit the restriction that

$$\mathrm{E}(v_t^{i,\ell}|\Delta \log y^1_{-\ell t}, \Delta \log y^2_{-\ell t}, \mathbf{1}_\ell) = 0,$$

where

$$\Delta \log y^i_{\ell t} = \gamma^i \Delta \log y^i_{-\ell t} + \eta^i_t + v^{i,\ell}_t.$$

In the present instance we do not exploit this last restriction. Our present approach is a more traditional way of approaching instrumental variables estimation, since traditional instrumental variables approaches typically place no restrictions on the relationship between the residual $v_t^{i,\ell}$ and the variables which appear in the estimating equation.

We would expect results from the present estimator to be less precisely estimated and the power of tests based on the present estimator to be of lower power than in Experiment 4.1, and indeed, this reduced precision and power is evident in Table 7. One can see from this table that almost all of the estimated coefficients are not significantly different from zero—there's a several-fold increase in the magnitude of the estimated standard errors associated with the coefficient estimates in the absence of the additional restriction (5).

Of course, precision isn't even really desirable if one's estimator isn't consistent, and one might be concerned that the point estimates reported in Table 7 are sufficiently different from those of our benchmark case that this might be evidence that the restrictions in (5) aren't satisified. Because the moment restrictions exploited by the present estimator and the estimator of Experiment 4.1 are nested, one way to test this is via a simple Hausman test of the equality of the coefficients across the two specifications.

It turns out that the coefficients in Table 7 (including the unreported error correction terms) are not significantly different from those reported in Table 3. Performing this test yields a statistic of $\chi^2_{440} = 247.30$, which has a p value of 1.0000.

4.6. **A Specification with Income Shares.** Sometimes a decomposition of GDP is used to motivate a regression which is somewhat similar to our equation of interest (2). In particular, let y_{it} denote the GDP of country i at time t. Suppose that income is derived from two sectors, so that

$$y_{it} = y^1_{it} + y^2_{it}.$$

Then

$$\Delta \log y_{it} \approx \theta^1_{it} \Delta \log y^1_{it} + \theta^2_{it} \Delta \log y^2_{it}, \tag{6}$$

where the approximation has to do with the log approximation to non-infinitesimal percentage changes.

Now, consider two alternative ways of estimating the parameters of (6). First, one could estimate

$$\Delta \log y_{it} = \beta_1 \Delta \log y^1_{it} + \beta_2 \Delta \log y^2_{it} + \epsilon_{it}; \tag{7}$$

second,

$$\Delta \log y_{it} = \gamma_1 \theta^1_{it} \Delta \log y^1_{it} + \gamma_2 \theta^2_{it} \Delta \log y^2_{it} + \epsilon_{it}. \tag{8}$$

Neither of these seem to be very interesting things to estimate. Least squares estimation of (8) will yield estimates of the parameters (γ_1, γ_2) of one, save for any approximation error in the approximation of changes in logs to growth rates which might be correlated with the sectoral growth rates. Estimating (7) on the other hand, will return estimates of the β_j which (again neglecting approximation error) simply equal to the average shares of the sectors.

Nevertheless, the difference between these two equations suggests a different specification for the estimation problem we're interested in. In particular, if we were to modify (2) so that the growth rates (changes in logs) of different sectors were premultiplied by lagged shares of those sectors, we'd obtain coefficient estimates which would be easier to interpret if there's much variation in shares across time or countries.

Accordingly, we compute lagged[2] shares of agricultural and non-agricultural sources of income across time and countries, and call these $\theta^j_{\ell t}$, $j = 1, 2$. We then re-specify our estimating equation as

$$\Delta \log c_t^{(\ell,q)} = \alpha^{(\ell,q)} + \beta^1 \theta^1_{\ell,t-1} \Delta \log y^1_{\ell t} + \beta^2 \theta^2_{\ell,t-1} \Delta \log y^2_{\ell t} + \epsilon_t^{(\ell,q)}. \tag{9}$$

	Border Instrument		
	Constant	Year Effects	Country Effects
Observations	2060	2060	2060
Agriculture			
Coefficient Est.	0.367**	0.409***	0.268**
Std. Errors	(0.115)	(0.117)	(0.120)
R^2	0.058	0.250	0.177
Non Agriculture			
Coefficient Est.	0.608***	0.424***	0.564***
Std. Errors	(0.088)	(0.098)	(0.102)
R^2	0.224	0.342	0.498

Table 8. First stage regression of the growth rate of agricultural and non-agricultural income on the average of neighboring countries' growth rates of these different sources of income. Income growth rates are weighted by the lagged share of agricultural and non-agricultural income. Different columns reflect different error-correction strategies. The first column includes only a constant; the second a collection of year-dummy variables; and the third a collection of country-fixed effects.

Another way to think of this is simply that our data on income growth is now weighted by lagged shares. Using these weighted data to replace the unweighted income data that appears in Experiment 4.1 yields results from our first stage which are reported in Table 8. Perhaps

[2]The unbalanced nature of our panel makes the use of the word "lagged" somewhat misleading. By lagged we don't necessarily mean data from the previous year, but rather from the previous year of available data; we ignore this problem in the notation below when we use the time subscript $t-1$.

surprisingly, the fit of this first-stage is actually somewhat improved relative to the specification in Experiment 4.1, with R^2 statistics improving somewhat for every specification.

Deciles	Agricultural Income Growth	Non-Agricultural Income Growth
Shares	0.226	0.774
Std. Errors	0.283	0.522
10%	6.151^{***}	-0.889^{*}
20%	3.926^{***}	1.044^{**}
30%	3.202^{***}	1.305^{**}
40%	2.795^{***}	1.450^{**}
50%	2.572^{***}	1.532^{**}
60%	2.425^{***}	1.609^{**}
70%	2.326^{***}	1.649^{**}
80%	2.167^{***}	1.663^{**}
90%	1.917^{***}	1.625^{**}
100%	0.437	1.571^{**}

Table 9. Second stage regression of the growth rate of decile expenditures on the aggregate growth rates of agricultural and non-agricultural income. The annotations *, **, and *** indicate significance of the corresponding coefficient estimate with levels of confidence corresponding to 90%, 95%, and 99%. In addition to the reported variables, the (second stage) estimating equation includes fixed effects for each country-decile, while the first stage includes year effects.

Results from the second stage regression are shown in Table 9. Using income growth variables weighted by shares changes the interpretation of the point estimates, of course. While in Experiment 4.1 the correct interpretation of the reported point estimates had to do with the elasticity of expenditures with respect to, say one percent growth in a particular sector, in Table 9 the correct interpretation is roughly that of an elasticity with respect to one percent growth in total GDP *due* to a particular sector. To be more precise, if all countries had unchanging, identical sectoral shares, then one could obtain the estimates in Table 9 simply by dividing the estimates in Table 3 by those shares.

This last way of thinking about the relationship between the estimates of Experiment 4.1 and the present estimates suggests that one

Figure 1. Plot of expenditure elasticities.

could draw some indirect inferences regarding the importance of variation in shares across time or countries by comparing the estimates of Table 3 divided by average shares to the estimates of Table 9. And as it happens, this exercise yields point estimates which are reasonably similar, suggesting that this source of variation is not key to understanding the relationship between expenditure and income growth across deciles.

Despite the apparent unimportance of variation in shares across time or countries for our estimation strategy, for some purposes the interpretation of elasticities in the weighted case may be preferred. Table 9 uses the hypotheses of homoskedasticity and independence to estimate standard errors. Table 10 improves upon this using the robust estimator of the covariance matrix of our estimates discussed in Experiment 4.2. As before, there's a substantial increase in the magnitude of the estimated standard errors (though generally by less than a factor of two) between the two tables. The extra information on shares, meanwhile, sharpens the precision of our estimates, particular as concerns the effects of agricultural income growth on expenditure growth, so

	Sectoral Income Growth		
Deciles	Agricultural	Non-Agricultural	Test
10%	6.151***	−0.889	90.273
	(0.414)	(0.826)	(0.000)
20%	3.926***	1.044	12.044
	(0.438)	(0.838)	(0.001)
30%	3.202***	1.305	5.814
	(0.466)	(1.032)	(0.016)
40%	2.795***	1.450**	4.496
	(0.378)	(0.656)	(0.034)
50%	2.572***	1.532**	4.430
	(0.426)	(0.717)	(0.035)
60%	2.425***	1.609*	1.137
	(0.421)	(0.904)	(0.286)
70%	2.326***	1.649*	0.844
	(0.503)	(0.856)	(0.358)
80%	2.167***	1.663**	0.580
	(0.390)	(0.728)	(0.446)
90%	1.917***	1.625**	0.176
	(0.599)	(0.808)	(0.675)
100%	0.437	1.571***	8.353
	(0.538)	(0.542)	(0.004)

Table 10. Second stage regression of the growth rate of decile expenditures on the aggregate growth rates of agricultural and non-agricultural income weighted by lagged shares. In the first two columns the annotations *, **, and *** indicate significance of the corresponding coefficient estimate with levels of confidence corresponding to 90%, 95%, and 99%. In addition to the reported variables, the (second stage) estimating equation includes fixed effects for each country-decile, while the first stage includes year effects. Reported standard errors are robust to arbitrary forms of heteroskedasticity, and arbitrary patterns of correlation across deciles. The final column reports chi^2 statistics associated with the null hypothesis that the coefficients reported in the first two columns are equal. Parenthetical numbers in the final column are p-values associated with this test of equality.

that all of the resulting point estimates are positive and highly significant, save for the very top decile. In contrast, the estimated elasticities associated with growth outside of agriculture are no longer significant for the bottom three deciles.

5. Conclusion

In this paper we've explored some different approaches toward estimating the effects of agricultural growth on expenditure growth and distribution. Our basic approach takes advantage of the fact that we have data on both aggregate rates of expenditure growth across countries, and on changes in the distribution of these expenditures across households.

We improve on much of the existing literature by not only providing evidence that sectoral GDP growth is endogenous (presumably it is jointly determined along with expenditures), but also devising an instrumental variables strategy to correct for this endogeneity, involving averaging over income growth rates for neighboring countries. We also improve on much of the literature by taking full advantage of the panel aspect of these data, a task which is considerably complicated by the extremely unbalanced nature of the panel.

We find that agricultural income growth has a particularly beneficial effect on expenditure group for the poorest households (in terms of expenditures), while the benefits of non-agricultural income growth are much more modest for households in lower deciles. Conversely, the benefits of agricultural income growth dissipate for households in higher expenditure deciles, while the benefits of non-agricultural income growth are increasing. The evidence presented here is generally consistent wth the view that while agricultural income growth is more effective at reducing poverty than is growth in other sectors.

These general results are robust to a variety of extensions and robustness checks. We experiment with a an estimator which allows for the estimation of covariance matrices which are robust to arbitrary forms of heteroskedasticity and correlation patterns across the different deciles within a country-year; this tends to increase our estimated standard errors, but our basic findings (and significance of key coefficients) is unaffected. We could adopt this strategy in a more wholesale fashion, but feel that caution is warranted, since the finite sample properties of this class of estimators is known to be sometimes quite poor (and our sample is very finite!).

We also estimate our main equation of interest using Ordinary Least Squares. This leads to significantly smaller estimated elasticities, allowing us (on the basis of a Hausman test) to reject the null hypothesis that income growth is exogenous. Despite this, the basic pattern of relatively large, positive elasticities of expenditure growth with respect to agricultural income growth for poorer deciles holds, even without correcting for the evident endogeneity of income growth.

We experiment with dividing the sample according to whether a country has high or low initial shares of agricultural growth, on the hypothesis that elasticities may be quite different for countries in which agriculture is relatively important. Here we seem to encounter once again the limitations of our dataset; our results are consistent with the view that households with higher initial shares of agriculture have higher elasticities, but the results also feature some wildly improbable magnitudes of the estimated elasticities for both the poorest and wealthiest of households, so we choose not to pursue this approach.

We also construct a conventional instrumental variables estimator which exploits only a subset of the restrictions we use in our benchmark estimator. Accordingly, this estimator gives rise to less powerful tests and less precise esimates. However, it permits us to test the additional conditions exploited by the benchmark estimator; we're unable to reject the null hypothesis that the benchmark conditions are satisfied, leaving us to prefer this more efficient estimator. Nonetheless, this more traditional instrumental variables approach serves to highlight the greatest weakness of our approach, which seems to lie in the somewhat ad hoc approach we've taken to the specification of various latent "error correction" terms such as year effects, country fixed effects, and similar. Our benchmark estimator takes advantage of a particular error correction strategy involving the use of fixed effects for country-decile pairs in a second-stage estimation, and year effects in a first stage. Results are sensitive to the changes in this strategy, reflecting the very limited data available (a total of 151 country-year changes are available in the data) and the large reductions in power which attend the estimation of large numbers of error correction terms.

Finally, we experiment with an alternative specification in which rates of growth of agricultural and non-agricultural income are weighted by the lagged shares of agriculture and its complement. This changes the interpretation of the estimated elasticities, so that one can make statements having to do with the elasticity of expenditures for a given decile with respect to total income growth due to a particular sector; we also present estimates of these elasticities using our robust covariance matrix estimator. These results indicate that a one percent increase in GDP due to agriculture result in a more than six percent increase in expenditure growth for the poorest decile, and indeed has a significantly disproportionate effect on expenditure growth for all but the top two expenditure deciles. Non-agricultural income growth, on the other hand, is disproportionately beneficial for the upper expenditure deciles, but has no significant effect on expenditure growth for households in the bottom 30 per cent of the expenditure distribution.

References

Amemiya, T. (1985). *Advanced Econometrics.* Cambridge, Massachusetts: Harvard University Press.

Arellano, M. (1987). Computing robust standard errors for within-groups estimators. *Oxford Bulletin of Economics and Statistics 49*(4), 431–34.

Chen, S. and M. Ravallion (2004). China's (uneven) progress against poverty. Policy Research Paper 3408, World Bank, Washington, D.C.

Davis, D. R. and P. Mishra (2007). Stolper-Samuelson is dead: And other crimes of both theory and data. In A. Harrison (Ed.), *Globalization and Poverty*, National Bureau of Economic Research. Chicago: University of Chicago Press.

Loayza, N. and C. Raddatz (2006, December). The composition of growth matters for poverty alleviation. Policy Research Working Paper 4077, World Bank.

Ravallion, M. and G. Datt (1996). How important to India's poor is the sectoral composition of economic growth? *The World Bank Economic Review 10*(1), 1–25.

Thorbecke, E. and H.-S. Jung (1996). A multiplier decomposition method to analyze poverty alleviation. *Journal of Development Economics 48*(2), 279–300.

Dept. of Agricultural & Resource Economics, University of California, Berkeley