Discussion Paper No. 2003/54

Externalities in Rural Development

Evidence for China

Martin Ravallion*

July 2003

Abstract

The paper tests for external effects of local economic activity on consumption and income growth at the farm household level using panel data from four provinces of post-reform rural China. The tests allow for nonstationary fixed effects in the consumption growth process. Evidence is found of geographic externalities, stemming from spillover effects of the level and composition of local economic activity and private returns to local human and physical infrastructure endowments. The results suggest an explanation for rural underdevelopment arising from underinvestment in certain externality-generating activities, of which agricultural development emerges as the most important.

Keywords: consumption growth, income growth, externalities, panel data, rural China
JEL classification: D91, R11, Q12

* The World Bank

This study has been prepared within the UNU/WIDER project on Spatial Disparities in Human Development, directed by Ravi Kanbur and Tony Venables.

UNU/WIDER gratefully acknowledges the financial contribution to the project by the Government of Sweden (Swedish International Development Cooperation Agency—Sida).

Acknowledgements

The data used here were kindly provided by China's National Bureau of Statistics, and I am grateful for the assistance and advice provided by NBS staff in Beijing and at various provincial and county offices. For help with setting up the panel data set I am grateful to Shaohua Chen and Qinghua Zhao and for help with the calculations reported here I am grateful to Jyotsna Jalan. The support of the World Bank's Research Committee and a Dutch Trust Fund is also gratefully acknowledged. For their comments, I am grateful to Vernon Henderson, Jyotsna Jalan, Peter Lanjouw, Forhad Shilpi, Dominique van de Walle, participants and discussants at the WIDER/Cornell/LSE conference on spatial inequality held at the London School of Economics in June 2002, the World Bank and anonymous referees for this work. These are the views of the author, and should not be attributed to the World Bank or any affiliated organization.

UNU World Institute for Development Economics Research (UNU/WIDER)
Katajanokanlaituri 6 B, 00160 Helsinki, Finland

Camera-ready typescript prepared by Lorraine Telfer-Taivainen at UNU/WIDER
Printed at UNU/WIDER, Helsinki

ISSN 1609-5774
ISBN 92-9190-494-5 (printed publication)
ISBN 92-9190-495-3 (internet publication)

1. Introduction

There is a long-standing view that externalities play an important causal role in economic development. Famously, Rosenstain-Rodan (1943) argued that the investment decisions made by one firm in a developing economy influenced the profitability of others, leading him to argue for international assistance for the industrialization of the lagging regions of Eastern and Southern Europe in the 1940s. More recently, the hypothesis that there are externalities through knowledge spillovers has been built into theoretical models of economic growth (notably Romer 1986; Lucas 1993). In the context of rural development in poor countries, similar ideas have motivated policy arguments that getting one activity going locally stimulates others, in a 'virtuous cycle' of growth; Mellor (1976) provided an influential statement of this hypothesis.[1] Hazell and Haggblade (1993) tested the hypothesis using district- and state-level data for India, and reported seemingly strong effects of agricultural growth on rural non-farm development.[2]

This paper explores the micro-empirical foundations of these arguments using household panel data for a developing rural economy. Some stylized facts about the setting will help motivate the subsequent analysis. One such fact is that in a poor rural economy, the income gains that are claimed to stem from linkage will be transmitted in large part through the farm household economy, which accounts for the bulk of rural economic activity in most developing countries. No doubt, spillover effects will also involve rural-based firms. However, it is plausible in this setting that any external impacts of local economic activity on income growth would be evident at the farm household level. A second stylized fact is that many farm households engage in multiple activities simultaneously, including non-farm activities. Casual observations do not suggest that it is commonly the case that a rural household is fully specialized in either farm or non-farm activities. Indeed, it has been argued that such income diversification is an important strategy by which rural households cope with uninsured risk (see, for example, Ellis 1998). There is a large literature pointing to the problems of incomplete credit and risk markets in underdeveloped rural economies (for an overview see Besley 1995).

It is not implausible that there are externalities in this setting. One way this happens is when farmers learn about new techniques of production from the experience of their neighbors; Feder and Slade (1985) provide survey evidence for northwest India that this is an important channel for knowledge diffusion amongst farmers. Foster and Rosenzweig (1995) find evidence of this type of learning externality in farm profitability from adopting

---

[1] Building on Mellor and Lele (1972). Much earlier still, Clarke (1940) had argued that higher agricultural productivity was a crucial precondition for industrialization.

[2] Also see Haggblade et al. (1989) and Haggblade et al. (2002). Lanjouw and Lanjouw (2001) provide a useful review of the arguments and evidence on the rural non-farm sector.

new seed varieties in India. Network effects in the marketing of agricultural products can also generate externalities: a farmer can benefit from the infrastructure already in place locally. Another possible source of externalities is the presence of local non-farm industries that encourage the acquisition of knowledge and skills that also benefit local farmers or non-farm enterprises at household level, possibly through knowledge sharing within households (Basu et al. 2002). In the case of China, it has been argued that higher output from the non-farm sector has brought external benefits to the traditional farm sector, through improved technologies and management (Sengupta and Lin 1995). Or a higher density of commercial enterprises may enhance the local tax base, allowing better local public goods, and so promoting higher growth for those not actually engaged in those enterprises. Alternatively, negative externalities might result when the expansion of one activity creates congestion, or otherwise crowds out, another activity. This can happen when there are local-level fixed factors of production (including environmental assets) that are shared across activities. For example, with imperfect credit markets leading to rationing of the available credit, an expansion in one activity may crowd out growth prospects in another. With restricted migration and wage stickiness, the same could happen with regard to labor.

If the patterns found in aggregate data reflect such externalities this would provide an important insight into the causal processes creating rural underdevelopment. That depends crucially on whether markets exist for the externalities.[3] That cannot be judged on a priori grounds. However, a complete set of such markets is not inherently plausible for the sorts of externalities discussed above. Knowledge spillovers or network effects do not lend themselves to the excludability properties needed for a market. (It would clearly be difficult to define and enforce property rights for such externalities.) So there must be a reasonable presumption that private decisionmakers will not typically take account of the external costs and benefits of their allocative decisions and so one will expect to see underinvestment in the activities that generate positive externalities, and over-investment in those that have negative externalities. The externalities then impede or distort rural development. On the other hand, if the underlying linkage effects are purely internal at the farm household level then their welfare and policy significance is greatly diminished.[4] Given the stylized facts summarized above, the averaging of purely internal effects within diversified farm household units could readily generate the appearance of externalities in economic activity in aggregate data when in fact none exist at the micro level. For example, given capital market imperfections, higher farm income for a given household may create the resources needed to finance a new non-farm activity. Farm and non-farm incomes may then co-move in a process that one might identify as intersectoral linkage in

[3] On the economic theory of markets for externalities, see Dasgupta and Heal (1979, chapter 3).

[4] It is often argued that the same is true if the externalities are 'pecuniary,' meaning that they are transmitted through prices. However, it is known that with incomplete markets, pecuniary externalities can still be a source of inefficiency (Greenwald and Stiglitz 1986; Hoff 1998 2000). The externality transmitted through prices could exacerbate the pre-existing inefficiency.

aggregate data even though there is no genuine externality involved. The causal connection is of course unclear, nor is it obvious that there would be any believable identification strategy.

The concern with geographic externalities goes beyond economic efficiency. It also raises concerns about horizontal equity. In particular, if the micro growth process involves such externalities then the economy will reward otherwise identical individuals differently depending on where they live. This may also help understand geographic dimensions of social unrest, as has been reported in China in the 1990s.[5] Motivated by these observations, the central question addressed in this paper is whether the signs of linkage amongst economic activities found in geographic data stem from externalities. From what we know about the features of a developing rural economy it is clear that one cannot conclude from the existing literature on linkages in rural development that externalities are present to any significant extent. The signs of linkage in geographically aggregated data could easily stem from a process in which there is in fact no interdependence amongst individual farm household units.

Testing for externalities poses a problem, even with micro panel data. Correlations between individual outcomes and geographic variables have been widely reported in the literature. However, as is well recognized, one cannot assume that the geographic placement of economic activity is exogenous at the micro level.[6] Placement in a given locality cannot be expected to be independent of the characteristics of the households that live there—no doubt including characteristics that are unobserved by the analyst. Persistent spatial concentrations of individuals with personal attributes that inhibit growth in their living standards, and lead to a worse assignment of geographic assets, can readily entail that the cross-sectional correlations often found in the data are entirely non-causal, with little or no bearing on development policy. All one is really picking up in the data is the fact that households who are poor in terms of some latent characteristic tend to be grouped together spatially and are less able to attract infrastructure and other geographically assigned resources.

To make this argument more concrete, consider the following example. In any rural economy, the quality of farmland is likely to be important to the productivity of current

---

[5] For example, an article on page one of *The New York Times* on 27 December 1995, stated: 'As China's economic miracle continues to leave millions behind, more and more Chinese are expressing anger over the economic disparities between the flourishing provinces of China's coastal plain and the impoverished inland, where 70-80 million people cannot feed or clothe themselves and hundreds of millions of others are only spectators to China's economic transformation.'

[6] For example, Foster and Rosenzweig (1996) report a significant coefficient on village placement of agricultural extension services in regressions for the adoption of high yielding varieties in micro data for India. As they point out, this cannot be considered a causal effect since the placement of extension services may depend on geographically-associated latent factors influencing adoption.

and past investments and hence economic growth. Land quality tends to be spatially correlated; the quality of one farmer's land is positively correlated with the quality of his neighbor's. However, land quality is rarely captured well even in quite comprehensive surveys. At the same time, one can expect that the composition of economic activity and the placement of rural infrastructure (irrigation, roads, etc.) will be influenced by land quality. In such circumstances, one can expect to find correlations between one farmer's income growth rate over time and the attributes of the area in which he lives, even controlling for observable characteristics of the farmer, such as his capital stock. That correlation might look like an externality, but it may simply be picking up the geographically associated latent heterogeneity in land quality.

The paper presents results of a test for geographic externalities through the composition of economic activity that is robust to such latent heterogeneity. Both household panel data and geographic data are clearly called for to have any hope of identifying geographic externalities in the growth process at the micro level. In modeling such data, one might turn to a standard panel data model with a time-invariant error component, as in (for example) the regressions for farm profits in Foster and Rosenzweig (1995). Allowing for latent heterogeneity in the household-level growth process will protect against spurious geographic effects due to time-invariant omitted variables. However, standard panel-data methods of eliminating the household-specific effect wipe out the time-invariant geographic variables of interest in this context, namely the initial composition of economic activity in the locality. Nor is it plausible that the latent heterogeneity in growth rates is time invariant; macroeconomic and geoclimatic conditions might well entail that the impact of these individual effects varies from year to year.

However, by simply relaxing the assumption that the fixed effect has a time-invariant impact one can estimate the effect of geographic differences in the observed initial level of economic activity on the micro growth process robustly to the latent heterogeneity. In particular, the analysis in this paper allows for nonstationary individual effects, following Holtz-Eakin et al. (1988) and Jalan and Ravallion (2002). The analysis combines geographic data on the composition of economic activity and infrastructure endowments with longitudinal micro observations of consumption and income growth by sector. The growth rate of household consumption is decomposed by income source to explore the income effects of geographic differences in the composition of economic activity and other geographic characteristics. This allows a reasonably flexible description of the patterns of externalities within and between sectors of the economy, as they affect the growth process. The following section outlines the econometric model. Section 3 describes the setting and data while section 4 presents the results. Section 5 summarizes the conclusions.

2 Econometric model

The aim is to test for external effects of the local composition of economic activity on the consumption growth process at the micro level. To provide a theoretical motivation for the empirical work, let us start with the standard assumption that households maximize an intertemporally additive utility integral, with common preferences. Since this is an agricultural setting, each household owns a production technology, and one can make the standard assumption that its output is a concave function of the household's own-capital. Add to this the (non-standard) assumption that output also depends non-separably on characteristics of the area of residence, including the composition of economic activity. If we further add restrictions on capital mobility, then marginal products of owncapital will not be equalized across farm households, but will differ geographically. The optimal rate of consumption growth will depend on the farm household's marginal product of own capital, which in turn depends on both the farm household's capital stock and its geographic characteristics, including the composition of economic activity locally.

The key feature of this model for the present purpose is that geographic externalities can influence consumption growth rates at the farm household level, through their effects on the productivity of private investment, given capital market imperfections. (The extreme case in which markets worked perfectly would imply that one had no power to explain the growth in consumption at the farm household level.) To derive a parametric model, one can assume constant discount and time preference rates and that the marginal product of own capital at the farm household level can be written as a linear function of observable household and geographic characteristics.[7] Thus, one can postulate the following parametric model of consumption growth for N households observed over T periods:

$$\Delta \ln C_{it} = \alpha + \beta X_{it} + \xi Z_i + \varepsilon_{it} \qquad (i=1,..,N;\ t=2,..,T) \qquad (1)$$

where C_{it} is consumption by household i at date t, $\Delta \ln C_{it}$ is the growth rate of consumption, X_{it} is a vector of time-varying explanatory variables, and Z_i is a vector of exogenous time-invariant explanatory variables including measures of the initial economic activity in the locality in which household i lives. (The properties of the error term, ε_{it}, are discussed below.) The variables in X_{it} and Z_i capture observable differences between households that influence the marginal product of their own-capital; these variables include indicators of the initial level (at time t-1) of own-capital plus relevant geographic variables, including the composition of economic activity locally.

Notice that in this model, one is testing for significant effects of the local composition of economic activity on the consumption growth rate at the household level, controlling for pre-determined household characteristics that directly influence the marginal product of own capital. This can be thought of as a key structural equation within a more complete model that also accounted for the evolution over time of the own-capital stock, which can then also be postulated as depending on the geographic variables. Intuitively, one might

---

[7] Alternatively one can allow for heterogeneity in discount rates or preferences, which can be taken to vary with the same characteristics. The interpretation in terms of productivity effects is then lost, since the geographic variables could in principle influence the intertemporal parameters.

expect lower rates of capital accumulation in areas for which the local composition of economic activity leads to lower productivity of private capital. Thus one can postulate a more complex multi-equation dynamic model that traces out these various direct and indirect effects of the composition of local economic activity on the evolution of household living standards. In this paper, the concern is solely with the key structural equation, which identifies the external effect of local economic activity on the productivity of own-capital, and hence consumption growth, given the level of own-capital.

The assumptions made about the error term in (1) are of course critical. One naturally wants to include a fixed error component that may well be correlated with the regressors of interest, as discussed in the introduction. The potential endogeneity of the explanatory variables in (1) is assumed to be fully captured by non-zero correlations with this error component. However, it is not assumed that the impact of the heterogeneity is necessarily constant over time. For example, some farmers are more productive than others in ways that cannot be captured in the data and this matters more in a bad agricultural year than a good one. Following Holtz-Eakin et al. (1988), the specification of the error term allows for nonstationarity in the impacts of the individual effects:

$$\varepsilon_{it} = \theta_t \omega_i + \mu_{it} \quad (2)$$

where μ_{it} is the i.i.d. random variable, with zero mean and variance σ_μ^2 , and ω_i is a time-invariant effect that is not orthogonal to the regressors, i.e., $E(\omega_i X_{it}) \neq 0$ and $E(\omega_i Z_i) \neq 0$, while μ_{it} is a white-noise innovation process, i.e., $E(\omega_i \mu_{it}) = 0$ and $E(Z_i \mu_{it}) = 0$. The assumed error structure in (2) facilitates quasi-differencing of the model in (1). Substituting equation (2) into (1) and lagging by one period one obtains:

$$\Delta \ln C_{it-1} = \alpha + \beta X_{it} + \xi Z_i + \theta_{t-1}\omega_i + \mu_{it-1} \quad (3)$$

Multiplying equation (3) by $r_t \equiv \theta_t / \theta_{t-1}$ and subtracting from equation (1), the quasi-differenced model for consumption growth is:

$$\Delta \ln C_{it} = \alpha(1 - r_t) + r_t \Delta \ln C_{it-1} + \beta(X_{it} - r_t X_{it-1}) + \xi(1 - r_t) Z_i + \mu_{it} - r_t \mu_{it-1} \quad (4)$$

It is evident from (4) that as long as $r_t \neq 1$ one can identify the impact of the time-invariant variables on the growth rate robustly to latent heterogeneity. The test described in Jalan and Ravallion (2002) (following Godfrey, 1988) is used to test the null hypothesis that $r_t = 1$ for all *t*. In estimating equation (4) one must allow for the fact that $\Delta \ln C_{it-1}$ is correlated with the error term, $\mu_{it} - r_t \mu_{it-1}$. One can estimate equation (4) by Generalized Method of Moments (GMM) using differences and/or levels of log consumptions lagged twice (or higher) as instruments for $\Delta \ln C_{it-1}$. (So one loses two observations over time in estimating equation 1.) The essential condition to justify this choice of instruments is that the error term in (4) is second-order serially independent, as implied by serial independence of μ_{it}. The Arellano and Bond (1991) second-order serial correlation test is performed, given that the consistency of the estimator for the quasi-differenced model depends on the assumption that the composite error term is second-order serially independent.[8] Note that there is some first-order serial correlation introduced in the model due to the quasi-differencing. This means that consumption lagged once is not a valid instrument.

Let us now see how the household-level impacts on consumption growth identified using the above model can be decomposed by income source. There are M-1 income sources and let Y_{jit} denote income from source j for household i at date t and (for notational convenience) let Y_{Mit} denote savings. From the identity:

$$C_{it} = \sum_{j=1}^{M} Y_{jit} \tag{5}$$

we have:

$$\Delta \ln C_{it} \cong \frac{\Delta C_{it}}{C_{it-1}} = \sum_{j=1}^{M} \frac{\Delta Y_{jit}}{C_{it-1}} \tag{6}$$

This motivates a decomposition of equation (4) as follows:

$$\begin{aligned} &\Delta Y_{jit} / C_{it-1} - r_t \Delta Y_{jit-1} / C_{it-2} = \alpha_j (1 - r_t) + \beta_j (X_{it} - r_t X_{it-1}) + \\ &\xi_j (1 - r_t) Z_i + \mu_{jit} - r_t \mu_{jit-1} \ (j = 1,..,M) \end{aligned} \tag{7}$$

Summing equation (7) over all j yields equation (4), with $\alpha = \sum \alpha_j$, $\beta = \sum \beta_j$ and $\xi = \sum \xi_j$. Notice that for consistency with aggregation, the r_t (t=1,..,T) parameters cannot vary by income source. To estimate (7), I replace the r_t parameters by their estimates from the consumption growth model to give:

$$\begin{aligned} &\Delta Y_{jit} / C_{it-1} - \hat{r}_t \Delta Y_{jit-1} / C_{it-2} = \alpha_j (1 - \hat{r}_t) + \beta_j (X_{it} - \hat{r}_t X_{it-1}) + \\ &\xi_j (1 - \hat{r}_t) Z_i + \mu_{jit} - \hat{r}_t \mu_{jit-1} \ (j = 1,..,M) \end{aligned} \tag{8}$$

Thus, provided the individual effect has a time varying impact, one can identify geographic effects by income sources, which are robust to latent (individual or geographic) heterogeneity.

3 Setting and data

China experienced a surge in rural non-farm activity in the 1980s, in the wake of country-wide economic reforms (Byrd and Qingsong 1990). An important element of this was the

[8] To test if the instruments are valid, the Arellano and Bond (1991) over-identification test is also used. Lack of second-order serial correlation and the non-rejection of the over-identification test support our choice of instruments. For further discussion see Jalan and Ravallion (2002).

emergence and rapid growth of Township and Village Enterprises (TVEs). The fact that growth in the number of non-farm enterprises was preceded by more rapid agricultural growth (following decollectivization starting in the late 1970s) is sometimes interpreted as evidence of a strong forward linkage from agriculture to non-farm rural development in the Chinese setting. For example, Jiacheng (1990) argues that agricultural growth provided the key pre-condition for the rapid expansion of non-farm economic activities in the 1980s. However, there are other interpretations in the literature; for example, Haiyan (1990) argues that, while the stimulus for non-farm rural enterprise development came from agriculture, it was a *negative* stimulus, not positive—that the expansion of rural non-farm enterprises was stimulated by low agricultural productivity in certain regions.

Anti-poverty policy in China has emphasized poor area development programs, which have traditionally emphasized the role of agriculture (Leading Group 1988; World Bank 1992). There has been debate in policy circles about this emphasis on agriculture, with some people arguing that non-farm enterprise development should be given priority instead. There has also been a debate about whether these programs are effective in longer-term poverty reduction, or are simply short-term palliatives (with out-migration from poor areas seen by some as the only long-term solution). In previous work using these data, evidence was found of dynamic income gains from the central and provincial poor area development programs, implying quite reasonable economic rates of return (Jalan and Ravallion 1998).

The following analysis uses household-level data from China's Rural Household Survey (RHS) done by the State Statistical Bureau (SSB). A panel of 5,600 farm households spanning 111 counties over the six-year period 1985-90 was formed for four contiguous provinces in southern China, namely Guangdong, Guangxi, Guizhou, and Yunnan. The latter three provinces form southwest China, widely regarded as one of the poorest regions in the country. Guangdong on the other hand is a relatively prosperous coastal region (surrounding Hong Kong). The RHS is a well-designed and executed survey of a random sample of households in rural China, with unusual effort made to reduce non-sampling errors (Chen and Ravallion 1996). Sampled households fill in a daily diary on expenditures and are visited on average every two weeks by an interviewer to check the diaries and collect other data relevant to incomes. There is also an elaborate system of cross-checking at the local level. The consumption and income data from such an intensive survey process are almost certainly more reliable than those obtained by the common cross-sectional surveys in which the data are based on recall at a single interview. For the six-year period 1985-90 the survey was also longitudinal, returning to the same households over time. While this was done for administrative convenience (since local SSB offices were set up in each sampled county), the panel can still be formed.[9]

[9] Constructing the panel from the annual RHS survey data proved to be more difficult than expected since the identifiers could not be relied upon. Fortunately, virtually ideal matching variables were available in the

The income aggregate includes imputed values of revenues from own production (net of costs) valued at actual local selling prices (rather than the planning prices used in the original data; see Chen and Ravallion 1996). The consumption data include imputed values of the consumption streams from the inventory of consumer durables. Poverty lines designed to represent the cost at each year and in each province of a fixed standard of living were used as deflators. These were based on a normative food bundle set by SSB, which assures that average nutritional requirements are met with a diet that is consistent with Chinese tastes. This food bundle is then valued at province-specific prices. The food component of the poverty line is augmented with an allowance for non-food goods, consistent with the non-food spending of those households whose food spending is no more than adequate to afford the food component of the poverty line.[10] Income sources are broken down as follows:

(i) Farm income: income from grain production and other farm crops.
(ii) Non-farm income type I: forestry, animal husbandry, fishery, gathering and hunting.
(iii) Non-farm income type II: handicrafts, industry, material processing, construction, transportation, productive labor service, commerce, catering trade, services.
(iv) Collective income: collective production, income from TVEs, collective welfare funds, collective prizes, other collective income.

In adopting this classification, I wanted to distinguish the types of land-based non-farm income sources that are often associated with farming (type I) from others (type II). My usage is not standard in this respect; it is more common in the literature to only refer to my 'type II' as the 'non-farm sector' (see, for example, Lanjouw and Lanjouw 2001). Of course, in a literal sense, my 'type I' is not farming. And, as we will see, these three sectors behave differently, making their separation of interest. In 1985, these four income sources accounted for 58.4 percent, 24.5 percent, 15.0 percent and 2.1 percent (respectively) of aggregate household income in the sample. Multiple sources for one household are common. Indeed, every one of the sampled households who had income from farming also recorded at least some income from a non-farm activity.

Collective income is the most problematic of the four categories. Although income gains from non-household non-farm enterprises are excluded from this analysis, the profits received from such enterprises by households are included under 'collective income'. However, the category accounts for only 2 percent of income. And it is likely that some of this comes from outside the county. One can be justifiably skeptical as to how well the

---

financial records, which gave both beginning and end of year balances. The relatively few ties by these criteria could easily be broken using demographic (including age) data.

[10] For further details on the poverty lines and the data set see Chen and Ravallion (1996).

following analysis will then be able to capture external effects on local non-household income growth.

Echoing the empirical literature on linkages, one finds positive correlations across counties between farm income per capita and non-farm income of type I above, though less so for type II. Table 1 gives the correlation coefficients in the time-means in the data set. There is very little correlation between the two types of non-farm income.

Table 1: Correlation coefficients in sample mean incomes across 102 counties

	Farm Income	Non-farm income I	Non-farm income II	Collective income
Farm income	1.0000			
Non-farm income I	0.3240	1.0000		
Non-farm income II	0.1134	0.0027	1.0000	
Collective income	0.4505	0.1125	0.2171	1.0000

Source: See text.

In estimating equation (8), I shall use two distinct types of data on the geographic composition of economic activity. The first uses the initial (1985) county mean of the income sources identified above. Initial values of the corresponding household variables are also included. This gives a conceptually clean representation of the four-by-four matrix of linkage effects. However, there is a potential concern that the explanatory variables are from the same survey-based data source. There are of course sampling errors in the county means, and possibly correlated measurement errors. For the second set of estimates, I draw instead on county administrative data. This has two advantages. Firstly, the data sources are then largely independent, relieving possible concerns about correlated measurement errors when using a common data source. Secondly, the county administrative data encompass the rural non-household sector, including TVEs. A disadvantage is that the available county data are less complete, which reduces the sample size to 4,800 (96 counties).

From the county data, one can identify three obvious indicators of the extent of development of local agriculture, namely irrigated land area, fertilizer usage and agricultural machinery usage. For the rural non-farm sector, I have used the county administrative data on the number of commercial enterprises in 1985 and the sector composition of gross product per capita at county level. The latter is broken down according to whether it is industry (distinguished according to whether the industrial enterprise is township, village or household-based), construction, transport or services. In this second model, controls are also added for geographic and household heterogeneity. The geographic variables at the county-level database include population density, average education levels, road density, health indicators, and schooling indicators. Dummy variables for the province are also included. A composite measure of household wealth can be constructed, comprising valuations of all fixed productive assets, cash, deposits, housing, grain stock, and consumer durables. Data are also used on agricultural inputs used, including landholding. These asset and farm input variables are time-varying, but are treated as endogenous, using lagged values as instruments. To allow for differences in the quality and quantity of family labor (given that labor markets are thin in this setting), initial education attainments and demographic characteristics are also included. The Appendix provides descriptive statistics.

4 Results

First the simpler model described above is estimated, in which consumption growth and its components by income source are regressed on the survey-based estimates of initial county mean income by source and initial own incomes. Table 2 gives the consumption growth regression (corresponding to equation 4), while Table 3 gives the decomposition by all four income sources (equation 8). (Saving is the residual, not estimated.) The diagnostic tests described in section 2 passed comfortably. (This was also true of the extended model, discussed later in this section.) The results in Tables 2 and 3 are for the full sample (n=5,600); the models were also estimated on the smaller sample for which county data are complete (as used in the extended specification below). The results were very similar between the two samples, suggesting that there is nothing particularly unusual about the countries with incomplete county data. I chose to use all the available data for Table 2 rather than to limit the sample to the counties included in the regressions in Table 3.

Consumption growth at the household level is significantly higher in counties with higher initial levels of farm income, non-farm income type I and collective income. The size and significance of the effect of differences in county-mean farm income are notable; the regression coefficient in Table 2 implies that a 100 Yuan per month increase in mean farm income in the county of residence (equivalent to one standard deviation, or about 60 percent of mean farm income) increases the consumption growth rate by 0.0195—about two percentage points per annum. In marked contrast to the county variables, higher *own* incomes tend to result in lower subsequent consumption growth. This pattern echoes the finding of Jalan and Ravallion (2002) that the micro consumption growth process tends to be convergent with respect to household characteristics (in that characteristics that tend to raise the current level of consumption lead to lower subsequent growth), but divergent with respect to geographic characteristics.

Turning to the decomposition of consumption growth by income source, the results in Table 3 indicate a significant within-sector external effect in all cases except collective income. Higher initial mean incomes from farming in the county of residence entail higher subsequent income gains from farming. This is also the case for type II non-farm incomes. For type I non-farm incomes however, one finds a negative external effect within the sector, suggestive of a crowding-out effect.

Table 2: Consumption growth regressed on county-mean incomes and own incomes

Consumption growth 1985-90	GMM estimates	
	Coefficient	t-ratio
Constant	-0.019034*	-3.631332
Coefficients on lagged consumption		
1987	-0.023637	-0.260700
1988	0.231193*	5.477698
1989	-0.036034	-0.974515
1990	0.192418*	4.036306
County mean household incomes by source, 1985		
Farm income	0.000195*	7.029119
Non-farm income I	6.77E-05	1.848970
Non-farm income II	6.10E-05	1.376225
Collective income	0.000148	1.925260
Household's own income by source, 1985		
Farm income	-5.07E-05*	-3.616862
Non-farm income I	-7.25E-05*	-4.473417
Non-farm income II	-7.47E-05*	-5.055279
Collective income	-2.25E-05	-0.795760

Note: * indicates significant at 1% level, two-tailed test; *n*=5,641 (111 counties).

Source: See Text.

Looking at the cross-sectoral linkages in Table 3, one finds no significant effects of initial non-farm income in the county on farm income gains at the household level. A significant positive effect of a higher initial level of farm incomes in the county on the growth of type I non-farm incomes is found, but not for type II. Non-farm incomes of type I in turn have positive effects on the growth of type II and collective incomes. However, higher collective incomes locally tend to attenuate growth in non-farm incomes of type II. For each of the four income growth regressions in Table 3, one can convincingly reject the null hypothesis that the four coefficients on the county-mean income sources are equal.[11] Thus, the composition of economic activity matters. Summing the external effect of a given income component horizontally in Table 3, it is plain that farming is the largest generator of external effects on the growth process at micro level. Including savings, the sum across all components is given by the coefficient in the consumption growth regression. The aggregate external effects on consumption are positive for all four income components, but farming is the largest (Table 2). Across the four income components (excluding savings), the aggregate impact of higher farm income per capita in a county on subsequent income

[11] Wald tests of the null hypothesis that the four coefficients on county-mean incomes in Table 3 are equal gave 44.9, 35.8, 71.7 and 19.5 respectively. For the consumption growth regression in Table 2 the Wald test gave 15.3. The test has a Chi-square distribution with four degrees of freedom, implying rejections of the null hypotheses at the 1 percent level or better.

gains (normalized by lagged consumption) is 0.058, as compared to –0.019, -0.012 and 0.009 for non-farm incomes types I and II and collective incomes respectively (Table 3).

And in all four cases, the bulk of the external effect is the 'own-effect' of higher income in a county on subsequent income growth. For farming, 99 percent of the external effect is the 'own-effect.'So, while we see signs of an external effect of higher initial agricultural development on the growth of non-farm incomes at the household level, it is quantitatively small. With regard to the effects of initial household income on income growth by source, we can see some strong signs of negative 'own-income' effects, suggestive of convergence due to diminishing marginal products of own capital. There are also negative effects of the other income components on farm income.

Table 4 and 5 give the extended specification, which exploits the county administrative records on output by source and indicators of local human and physical infrastructure. The regression in Table 4 indicates significant positive effects of initial agricultural development on consumption growth at the household level. We also see significant positive effects of a higher density of commercial enterprises. Higher mean industrial output at village level is also associated with higher subsequent consumption growth at the farm household level, though this is not true when the enterprises are township- or household-based. On the other hand, there are indications of negative external effects of construction and transport sector development. Some of the other geographic controls are suggestive of positive externalities from better local endowments of human and physical infrastructure; in particular, higher levels of literacy locally and higher road density promote higher consumption growth at household level.

The indications of geographic externalities are also evident in the decomposition by income source (Table 5). Echoing the results of Table 3, here too one finds strong indications that areas with more land and more developed agriculture tend to experience higher subsequent farm income gains; this effect is particularly strong for fertilizer usage, which is probably the best indicator in these data of the adoption of modern agricultural techniques. Cross-effects of initial agricultural development on non-farm incomes are evident, although (in contrast to Table 3) the effects are now evident for both type I and type II non-farm incomes. However, unlike Table 3, one now finds strong positive effects of the density of non-farm commercial development and industrial output on farm incomes.

Table 3: Decomposition of growth by income source

Income change 1985-90, normalized by initial consumption	Farm income		Non-farm income I		Non-farm income II		Collective income	
	Coefficient	t- ratio	Coefficient	t- ratio	Coefficient	t- ratio	Coefficient	t- ratio
Constant	-1.042776	-0.606657	0.007246	1.581473	-0.001584	-0.304262	-0.006296*	-2.776274
County mean household incomes by source, 1985								
Farm income	0.058360*	5.582170	9.02E-05*	3.587523	-1.43E-05	-0.527382	4.85E-06	0.405785
Non-farm income 1	-0.019292	-1.864275	-7.72E-05*	-2.922507	9.23E-05*	2.627234	7.64E-05*	4.199318
Non-farm income 2	-0.012158	-0.836240	-2.46E-05	-0.722362	0.000358*	7.216553	1.95E-06	0.156793
Collective income	0.009052	0.365210	7.86E-06	0.104518	-0.000232	-2.364964	8.38E-05	1.705881
Household's own income by source, 1985								
Farm income	-0.065339*	-7.964560	-2.23E-05	-2.032218	-2.72E-05	-2.032737	9.13E-07	0.225140
Non-farm income 1	-0.009548	-2.082792	-8.46E-05*	-5.212832	8.70E-07	0.069791	-1.86E-05*	-3.288436
Non-farm income 2	-0.005469	-1.357622	4.01E-06	0.394638	-4.41E-05	-1.707732	1.76E-06	0.414443
Collective income	-0.024780*	-2.654654	-1.04E-05	-0.393923	1.42E-05	0.485131	-0.000132*	-5.666340
J statistic	0.073149		0.037610		0.020013		0.005590	

Notes: * indicates significant at 1% level, two-tailed test; *n*=5,641 (111 counties).

Source: See text.

14

Table 4: Consumption growth model using geographic data from county administrative records

	Coefficient	t-Statistic
Constant	-0.328076*	-3.938664
Coefficients on lagged consumption		
1987	-0.563094*	-5.580720
1988	0.226777*	6.313155
1989	-0.031837	-0.878866
1990	0.264715*	6.118230
Economic activity at county level		
(a) Farm		
Cultivated area per 10,000 persons	0.003075*	3.424595
Fertilizer used per cultivated area	0.004131*	7.433107
Farm machinery used per cultivated area	0.000368*	2.651082
(b) Non-farm		
Number of commercial enterprises in county per 10,000 population	0.000220*	2.768617
Rural industry gross product per 10,000: township enterprises	-6.63E-05	-1.759901
Rural industry gross product per 10,000 persons: village enterprises	0.000415*	3.729650
Rural industry gross product per 10,000 persons: household enterprises	-1.77E-05	-0.173829
Rural construction gross product per 10,000 persons	-0.000154	-2.063245
Rural transportation gross product per 10,000 persons	-0.000509*	-3.639974
Rural gross product from services per 10,000 persons	0.000169	0.715551
Other geographic controls		
Guangdong (dummy)	0.037373*	4.338988
Guangxi (dummy)	0.022666*	4.345667
Yunnan (dummy)	-0.005237	-0.869316
Revolutionary base area (dummy)	0.050238*	3.248796
Border area (dummy)	0.002216	0.563537
Coastal area (dummy)	-0.012471	-1.278915
Minority area (dummy)	-0.012457*	-3.714323
Mountainous area (dummy)	-0.015838*	-4.452355
Plains (dummy)	0.005659	1.459167
Population density (log)	0.021519	2.480439
Proportion of illiterates in 15^+ population	-0.000322	-1.866172
Infant mortality rate	-0.000147	-1.296671
Medical personnel per capita	0.000584	1.988495
Kilometers of roads per capita	0.000455*	3.185796
Proportion of population living in urban areas	-0.097467*	-3.199404
Household variables		
Expenditure on agricultural inputs per cultivated area	-0.001911*	-7.161740
Fixed productive assets per capita	-1.27E-05	-0.883144

Cultivated land per capita	-0.008748	-1.802922
Household size (log)	0.056994*	8.967627
Age of household head	0.002086*	2.617436
Age^2 of household head	-2.57E-05*	-2.899381
Proportion of adults in the household who are illiterate	0.007032	1.125765
Proportion of adults with primary school education	7.77E-06	0.001468
Proportion of children 6-11 years	0.013395	1.377193
Proportion of children 12-14 years	0.032215*	2.502249
Proportion of children 15-17 years	0.002467	0.158605
Proportion of children with primary school education	-0.002868	-0.736394
Proportion of children with secondary school education	0.020066	2.002172
Whether a household member works in the state sector (dummy)	-0.001098	-0.147539
Proportion of 60^+ members in the household	0.002312	0.187774

Notes: * indicates significant at 1% level, two-tailed test; *n*=4,778 (96 counties).
Source: See text.

By allowing us to break up non-farm incomes by sector (industry, construction, transport and services) the regressions using the county administrative data in Table 5 reveal that the more aggregate effects identified in Table 3 disguise some potentially important differences between sub-sectors. Indeed, while there are generally positive external effects of local industrial development, we see signs of *negative* external effects on farm and non-farm income growth of greater local activity in the transport and service sub-sectors. (Notice that the transport income effect is probably not picking up an effect of transport infrastructure, since I am controlling for road density.) It appears that these sectors are competing with household-level farm and non-farm activities for limited local resources that enhance the productivity of private investment and hence income growth at the farm household level.

Higher cultivated area per person in a county has a significant positive effect on the growth of non-farm type I incomes, but the (positive) effect on type II is barely significant at the 10 percent level. These findings lead one to question the claims sometimes made (in the case of China, see Haiyan, 1990) that a shortage of cultivated land in an area was an inducement to non-farm activities. One finds the opposite to be the case for non-farm activities by the household, though there is a sign of this effect on collective income (which here includes income from enterprises). Higher fertilizer usage also has an external effect on both types of non-farm income growth, though the dominant external effect is on farm incomes. The extended models in Table 5 also point to some diverse and in some cases surprising impacts across income sources. The positive effect of higher population density on consumption growth (Table 4) appears to be transmitted entirely through non-farm type I income growth. The effect of road density appears to be largely through higher farm incomes. Lower infant mortality (as an indicator of health care more generally) appears to have high returns to non-farm (type II) income growth. Higher basic education appears to spillover more into farming.

Table 5: Decomposition by income source

Income change 1985-90, normalized by initial consumption	Farm income		Non-farm income I		Non-farm income II		Collective income	
	Coefficient	t-ratio	Coefficient	t- ratio	Coefficient	t- ratio	Coefficient	t- ratio
Constant	-0.037285	-0.551930	-0.317851	-5.899826	0.017037	0.312113	0.002492	0.213991
Economic activity at county level								
(a) Farm								
Cultivated area per 10,000 persons	0.001425	1.975995	0.004668*	7.661727	0.00096	1.694963	-5.92E-05	-0.494620
Fertilizer used per cultivated area (x100)	0.3298*	7.450607	0.1553*	4.623649	0.0775*	2.335067	0.000321	0.042174
Farm machinery used per cultivated area (x100)	0.0137	0.881606	0.00559	0.539991	-0.0262*	-2.185012	0.00757*	2.829688
(b) Non-farm								
Number of commercial enterprises per 10,000 population	-6.40E-05	-1.012804	1.02E-05	0.192262	0.000256*	4.728632	-1.06E-06	-0.084385
Rural industry gross product per 10,000: township enterprises	1.53E-05	0.472816	-7.68E-05*	-3.695281	-5.09E-05	-2.126782	3.83E-05*	3.920372
Rural industry gross product per 10,000 persons: village enterprises	0.000128	1.339266	-1.44E-05	-0.280215	0.000355*	4.522848	-2.83E-05	-1.738851
Rural industry gross product per 10,000 persons: enterprises owned by households	0.000207*	2.549566	0.000188*	3.230548	-9.72E-05	-1.487319	-3.74E-05	-1.336554
Rural construction gross product per 10,000 persons	-1.16E-06	-0.017589	5.40E-05	1.257481	-2.94E-05	-0.546224	1.46E-06	0.136096
Rural transportation gross product per 10,000 persons	-0.000225	-1.893099	-0.000234*	-2.760754	-0.000240	-2.109118	-3.10E-05	-1.163832
Rural gross product from services per 10,000 persons	-0.000870*	-4.433674	0.000235	1.700295	3.25E-05	0.212639	-4.47E-05	-1.095567
Other geographic controls								
Guangdong (dummy)	0.040192*	5.578835	-0.039672*	-7.215861	-0.001470	-0.261715	0.001851	1.382536

17

Income change 1985-90, normalized by initial consumption	Farm income Coefficient	t-ratio	Non-farm income I Coefficient	t- ratio	Non-farm income II Coefficient	t- ratio	Collective income Coefficient	t- ratio
Guangxi (dummy)	0.007369	1.729346	0.006060	1.656530	0.000142	1.066986	0.000817	1.080041
Yunnan (dummy)	-0.008800	-1.728590	-0.003452	-0.851059	-0.000270*	-3.355940	0.000545	0.646179
Revolutionary base area (dummy)	0.051141*	3.953926	-0.005317	-0.896477	-0.000388*	-1.856168	0.000267	0.172994
Border area (dummy)	0.015034*	4.753477	-0.006759	-2.516487	6.47E-05	0.512528	-0.000337	-0.645808
Coastal area (dummy)	-0.050290*	-5.715417	0.001306	0.172308	0.007192	0.300917	0.002023	0.869084
Minority area (dummy) (x100)	-0.2865	-1.020337	-0.6569*	-2.915940	-0.01381	-0.598987	0.0167	0.369802
Mountainous area (dummy)	-0.019013*	-6.629943	0.004989	2.179428	0.005246	2.187413	0.000247	0.478782
Plains (dummy)	0.003090	0.882888	0.009975*	3.684431	0.002272	0.751968	-0.000332	-0.527546
Population density (log)	0.003167	0.445301	0.026073*	4.853929	-0.001470	-0.261715	-0.000364	-0.297651
Prop of illiterates in 15^{+} population (x100)	-0.0397*	-2.810640	0.0183	1.371223	0.0142	1.066986	0.00253	0.982911
Infant mortality rate	-3.89E-05	-0.437105	-0.000126	-1.605491	-0.000270*	-3.355940	-3.62E-05	-2.187786
Medical staff per capita (x100)	0.0368	1.435438	-0.00388	-0.148288	-0.0388*	-1.856168	0.00424	1.070008
Kilometers of roads per capita (x100)	0.0678*	5.693957	-0.0100	-0.985167	0.00647	0.512528	0.00121	0.663914
Proportion of population living in urban areas	-0.082497*	-3.331008	0.039684	1.939799	0.007192	0.300917	-0.004783	-0.796981
Household-level variables								
Expenditure on agricultural inputs per cultivated area (x100)	-0.1788*	-9.474159	-0.00532	-0.578969	-0.0199	-2.058113	9.48E-06	0.435807
Fixed productive assets per capita (x100)	-0.000515	-0.512903	8.22^{E}-05	0.074993	0.00457*	2.566003	-0.000204	-0.867348
Cultivate land per capita	-0.008281	-1.818687	-0.007585	-2.438985	-0.008547*	-3.224610	-0.000337	-0.511164
Household size (log)	0.012321*	2.596482	0.014103*	2.919817	0.002724	0.615920	-0.000927	-1.012190
Age of household head	0.000470	0.718755	0.000379	0.635091	-0.000143	-0.251665	5.26E-05	0.476080
Age^{2} of household head (x100)	-0.000731	-0.999952	-0.000355	-0.522368	2.45E-06	0.385523	-0.0001	-0.775087
Prop of adults in household who are illiterate	-0.002009	-0.373594	0.000758	0.185174	-0.001015	-0.229459	0.002438	2.292417
Prop of adults in household with primary school education	-0.002942	-0.667418	0.004343	1.223840	-0.005701	-1.464877	0.001718	1.771110
Prop of children in the household ages 6-11 years	0.005990	0.711307	0.009385	1.432443	-0.006400	-0.877534	0.001215	0.583170

18

Prop of children in household ages 12-14 years	0.004234	0.378313	0.012715	1.410109	0.012735	1.341365	0.003008	0.917733
Prop of children in household ages 15-17 years	0.002616	0.209421	-0.005501	-0.529853	0.018894	1.796201	0.005167	1.699518
Prop of children with prim school education (x100)	-0.0409	-0.121831	-0.3429	-1.283922	0.001907	0.612289	0.0189	0.220488
Prop of children with secondary school education	-0.001224	-0.149337	0.010930	2.011010	-0.007282	-1.247138	0.002381	1.411268
Household member works in state sector (dummy)	-0.018599*	-3.088751	-0.004086	-0.805931	-0.003461	-0.765217	-0.000913	-0.493164
Proportion of 60^+ members in the household	0.002762	0.261581	-0.005292	-0.646863	0.001195	0.151179	-0.000654	-0.363943

Notes: * indicates significant at 1% level, two-tailed test; *n*=4,778 (96 counties).

Source: See text.

19

5 Conclusions

The literature on linkages in rural development has largely ignored what is surely the most relevant question for policy: do the signs of linkage found in geographic data reflect externalities at the level of the individual decisionmaker? The data and methods used in past empirical work cannot distinguish externalities from other factors far more benign from a policy point of view. Yet the implications for understanding rural underdevelopment, and the implications for policy, depend crucially on whether the aggregate appearance of intersectoral linkage in rural development stems from externalities at the micro level.

The paper has offered a test that can identify any genuine linkage externalities, and can also test for micro effects on the growth process of differing geographic endowments of human and physical infrastructure. The paper has implemented the test using data for rural China during the post-reform period of farm and (particularly) non-farm rural development. The aim has been to describe the patterns of linkage in a way that is robust to latent heterogeneity. Like any description, the results beg many questions. In particular, the analysis has thrown little light on the precise sources of external effects. Are we seeing the effects of knowledge spillovers, or something else such as network externalities or pecuniary externalities?

The results do suggest that the level and composition of local economic activity has non-negligible impacts on consumption and income growth at the farm household level. There are significant positive effects of the level of local economic activity in a given sector on income growth from that income source. And there are a number of significant sectoral cross-effects, notably from farming to those categories of non-farm activities that tend naturally to be more linked to agriculture (forestry, animal husbandry, fishery), but also between the latter type of non-farm activity and other types (handicrafts, industry, processing, transportation etc.). Thus there is a direct link from the initial level of agricultural development to the first type of non-farm activities and a more indirect link to the second. There is less sign of the reverse linkage—from initial level of non-farm economic activity to growth in farm incomes. And there are indications of negative external effects from some non-farm activities, notably involving non-industrial subsectors (construction and transport). While I do find significant cross-sector effects, they are dwarfed by the within-sector effects. The composition (as well as the level) of local economic activity matters, and the sector that clearly matters most quantitatively is agriculture.

The results of this paper suggest that there are externalities at the farm household level underlying the signs of linkage found in more geographically aggregated data. Under the paper's identifying assumptions, the linkages found can be interpreted as genuine externalities, suggesting that private agents in this economy are not going to take account of all the potential income gains from their actions. Thus these results offer an explanation for rural under-development, arising from underinvestment in externality generating activities, notably agriculture and (to a lesser extent) certain non-farm activities. By the same token, the results offer a micro-empirical foundation for the long-standing, but poorly validated, claims in the literature about the potential for 'virtuous cycles' whereby a well-targeted external growth stimulus in a poor area can generate positive and more widely diffused income gains over time.

Thus these results offer support for the types of poor-area development programs that have been pursued by the Government of China since the mid-1980s. The emphasis that these programs have given to agricultural development is consistent with this paper's findings that agriculture is the key externality-generating sector of the Chinese rural economy. Of course, the detailed design of such programs is crucial, and this is not something that the results of this paper can throw much light on. However, the present results also point to the importance of local endowments of human and physical infrastructure to the micro-growth process. When combined with data on the costs to the government's budget of alternative interventions, these empirical results will hopefully also help inform public choices on how best to balance agricultural development initiatives with infrastructure development, so as to assure maximum growth of living standards in poor areas.

References

Arellano, M. and S. Bond (1991) 'Some Tests of Specification for Panel Data: Monte-Carlo Evidence and an Application to Employment Equation', *Review of Economic Studies* 58:277-98.

Basu, K., A. Narayan and M. Ravallion (2002) 'Is Literacy Shared within Households?', *LaborEconomics* 8:649-65.

Besley, T. (1995) 'Savings, Credit and Insurance', in J. Behrman and T.N. Srinivasan (eds) *Handbook of Development Economics* Vol. 3, North-Holland: Amsterdam.

Byrd, W. and L. Qingspong (1990) 'China's Rural Industrialization', in W.A. Byrd and L. Qingspong (eds) *China's Rural Industry: Structure, Development and Reform*, Oxford University Press: New York.

Chen, S. and M. Ravallion (1996) 'Data in Transition: Assessing Rural Living Standards in Southern China', *China Economic Review* 7:23-56.

Clarke, C. (1940) *The Conditions of Economic Progress*, Macmillan: London.

Dasgupta, P.S. and G.M. Heal (1979) *Economic Theory and Exhaustible Resources.* Cambridge University Press: Cambridge.

Ellis, F. (1998) 'Household Strategies and Rural Livelihood Diversification', *Journal of Development Studies* 35:1-38.

Feder G. and R. Slade (1985) 'The Role of Public Policy in the Diffusion of Improved Agricultural Technology', *American Journal of Agricultural Economics* 67:423-28.

Foster, A.D. and M.R. Rosenzweig (1995) 'Learning by Doing and Learning from Others: Human Capital and Technical Change in Agriculture', *Journal of Political Economy,* 103(6)1176-209.

Foster, A.D. and M.R. Rosenzweig (1996) 'Technical Change and Human-Capital Returns and Investments: Evidence from the Green Revolution', *American EconomicReview* 86:931-53.

Godfrey, L.G. (1988) *Misspecification Tests in Econometrics*, Cambridge University Press, Cambridge.

Greenwald, B. and J.E. Stiglitz (1986) 'Externalities in Economies with Imperfect Information and Incomplete Markets', *Quarterly Journal of Economics* 101(2):229-64.

Haggblade, S., P. Hazell and J. Brown (1989) 'Farm-Non-farm Linkages in Rural Sub-Saharan Africa', *World Development* 17(8)1173-201.

Haggblade, S., P. Hazell and T. Reardon (2002) 'Strategies for Stimulating Poverty-Alleviating Growth in the Rural Non-farm Economy in Developing Countries' mimeo, International Food Policy Research Institute: Washington DC.

Haiyan, Du (1990) 'Causes of Rapid Rural Industrial Development', in W.A. Byrd and L. Qingspong (eds) *China's Rural Industry: Structure, Development and Reform*, Oxford University Press: New York.

Hazell, P. and S. Haggblade (1993) 'Farm-Non-farm Growth Linkages and the Welfare of the Poor', in M. Lipton and J. van der Gaag (eds.) *Including the Poor,* World Bank: Washington DC.

Hoff, K. (1998) 'Adverse Selection and Institutional Adaptation', mimeo, Development Research Group, World Bank: Washington DC.

Hoff, K. (2000) 'Beyond Rosenstain-Rodan: The Modern Theory of Underdevelopment Traps', paper Presented at the Annual Bank Conference on Development Economics, World Bank: Washington DC.

Holtz-Eakin, D., W. Newey and H. Rosen (1988) 'Estimating Vector Autoregressions with Panel Data', *Econometrica* 56:1371-95.

Jalan, J. and M. Ravallion (1998) 'Are There Dynamic Gains from a Poor-Area Development Program?', *Journal of Public Economics* 67:65-85.

Jalan, J. and M. Ravallion (2002) 'Geographic Poverty Traps? A Micro Model of Consumption Growth in Rural China', *Journal of Applied Econometrics* 17(4):329-46.

Jiacheng, H. (1990) ‘Development Issues and Policy Choices’, in W.A. Byrd and L. Qingspong (eds) *China’s Rural Industry: Structure, Development and Reform*, Oxford University Press: New York.

Lanjouw, J.O. and P. Lanjouw (2001) ‘The Rural Non-farm Sector: Issues and Evidence from Developing Countries’, *Agricultural Economics* 26:1-23.

Leading Group (1988) *Outlines of Economic Development in China’s Poor Areas*, Office of the Leading Group of Economic Development in Poor Areas Under the State Council, Agricultural Publishing House: Beijing.

Lucas, R.E. (1993) ‘Making a Miracle’, *Econometrica* 61:251-72.

Mellor, J.W. (1976) *The New Economics of Growth: A Strategy for India and the Developing World*, Cornell University Press: Ithaca.

Mellor, J.W. and U. Lele (1972) ‘Growth Linkages of the New Grain Technologies’, *Indian Journal of Agricultural Economics* 18(1):35-55.

Romer, P.M. (1986) ‘Increasing Returns and Long-Run Growth’, *Journal of Political Economy* 94:1,002-37.

Rosenstain-Rodan, P. (1943) ‘Problems of Industrialization of Eastern and Southeastern Europe’, *Economic Journal* 53:202-11.

Sengupta, J.L. and B.Q. Lin (1995) ‘Recent Rural Growth in China: The Performance of the Rural Small-Scale Enterprises’, mimeo, University of California: Santa Barbara CA.

World Bank (1992) China: Strategies for Reducing Poverty in the 1990s, World Bank, Washington DC.

Appendix: descriptive statistics

	Mean	St. deviation
Dependent variables		
Average growth rate of consumption, 1986-90	0.0042	0.0777
Farm income: mean change as a proportion of lagged consumption, 1986-90	-0.0065	0.0687
Non-farm income I: mean change as a proportion of lagged consumption, 1986-90	0.0027	0.0688
Non-farm income II: mean change as a proportion of lagged consumption, 1986-90	0.0157	0.0755
Collective income: mean change as a proportion of lagged consumption, 1986-90	-0.0005	0.0244
Economic activity at county level		
Farm income, 1985 (Yuan/person/month)	161.0459	100.846
Non-farm income I, 1985 (Yuan/person/month)	92.9326	98.303
Non-farm income II, 1985 (Yuan/person/month)	62.9430	100.405
Collective income, 1985 (Yuan/person/month)	9.4162	40.971
Fertilizers used per cultivated area (tones per sq.km)	11.5402	6.6497
Farm machinery used per capita (horsepower)[a]	151.7879	110.2427
Cultivated area per 10,000 persons (sq km)	13.0447	3.2518
Number of commercial enterprises per 10,000 population	52.5922	22.003
Rural industry gross product per 10,000: enterprises in townships (central administrative villages)	32.7465	132.874
Rural industry gross product per 10,000 persons: enterprises in villages	16.2585	45.475
Rural industry gross product per 10,000 persons: enterprises owned by households	27.5416	33.049
Rural construction gross product per 10,000 persons	32.5597	42.9291
Rural transportation gross product per 10,000 persons	13.3423	0.9594
Rural gross product from services per 10,000 persons	22.6664	23.121
Other geographic variables		
Proportion of sample in Guangdong	0.1618	0.3683
Proportion of sample in Guangxi	0.3414	0.4742
Proportion of sample in Yunnan	0.2285	0.4199
Proportion living in a revolutionary base area	0.0191	0.1367
Proportion of counties sharing a border with a foreign country	0.1712	0.3767
Proportion of villages located on the coast	0.0316	0.1749
Proportion of villages in with an ethnic minority concentration	0.2978	0.4573

Proportion of villages that have a mountainous terrain	0.45563	0.498
Proportion of villages located in the plains	0.2292	0.4203
Population density (log)	8.20602	0.3929
Proportion of illiterates in the 15^+ population (%)	36.9547	16.0225
Infant mortality rate (per 1,000 live births)	43.24006	23.8535
Medical personnel per 10,000 persons	7.816894	5.0388
Kilometers of roads per 10,000 persons	14.7122	10.9721
Proportion of population living in the urban areas	0.0907	0.0548
Household level variables		
Expenditure on agricultural inputs (fertilizers & pesticides) per cultivated area (Yuan per mu)[a]	29.224	47.9954
Fixed productive assets per capita (Yuan per capita)[a]	129.8417	150.8919
Cultivated land per capita (mu per capita)[a]	1.2591	0.7802
Household size (log)	1.7086	0.3508
Age of the household head	41.8262	11.3887
Age^2 of the household head	1879.114	1015.252
Proportion of adults in the household who are illiterate	0.33876	0.2932
Proportion of adults with primary school education	0.3787	0.3074
Proportion of children 6-11 years	0.1199	0.1415
Proportion of children 12-14 years	0.0845	0.1071
Proportion of children 15-17 years	0.06796	0.0988
Proportion of children with primary school education	0.2780	0.3689
Proportion of children with secondary school education	0.0484	0.1709
Proportion of a household members working in the state sector	0.0421	0.2008
Proportion of 60^+ household members	0.06270	0.1222
Number of households		4,778
Number of counties		96

Notes: 1 mu = 0.000667 km^2; 'a' indicates time-varying variables.

Source: See text.