External Validity in a Stochastic World: Evidence from Low-Income Countries

Mark Rosenzweig, Yale University

Christopher Udry, Northwestern University

June 2018

Abstract

We examine empirically the generalizability of internally valid micro estimates of causal effects in a fixed population over time when that population is subject to aggregate shocks. Using panel data we show that the returns to investments in agriculture in India and Ghana, small and medium non-farm enterprises in Sri Lanka, and schooling in Indonesia fluctuate significantly across time periods. We show how the returns to these investments interact with specific, measurable and economically-relevant aggregate shocks, focusing on rainfall and price fluctuations. We also obtain lower-bound estimates of confidence intervals of the returns based on estimates of the parameters of the distributions of rainfall shocks in our two agricultural samples. We find that even these lower-bound confidence intervals are substantially wider than those based solely on sampling error that are commonly provided in studies, most of which are based on single-year samples. We also find that cross-sectional variation in rainfall cannot be confidently used to replicate within-population rainfall variability. Based on our findings, we discuss methods for incorporating information on external shocks into evaluations of the returns to policy.

A large proportion of empirical work in economics is concerned with estimating causal effects. These causal estimates are often used as the basis for policy recommendations and sometimes even policy initiatives. As empirical methods improve, most markedly with the rapid adoption of RCTs, it is appropriate to shift our focus of concern from issues of internal validity (e.g., identification) towards the external validity of estimates of causal effects. What can we learn from internally valid estimates of a causal effect of a treatment or action in a particular population in a specific set of circumstances about the effect of that same treatment or action in other populations and circumstances?

A common feature of studies that focus on causal estimates is that the estimates are obtained at one point in time. A small set of examples (that are either influential and/or by one of us) includes Duflo *et al*. (2011), Banerjee and Duflo (2008), Banerjee *et al*. (2013), Bloom *et al*. (2013), de Mel *et al*. (2008, 2009), Foster and Rosenzweig (1995), Hanna *et al*. (2014), Karlan *et al*. (2013), Fafchamps *et al*. (2011), Udry and Anagol (2006), Suri (2011). A potential shortcoming of such studies is that the influence of aggregate shocks cannot be identified. If these shocks importantly affect the estimates obtained, these single-year estimates are valid only for the specific time-period in which they are obtained, potentially severely limiting their external validity to the extent that aggregate shocks are highly variable.

While not all studies have ignored changing responses over time, papers that exploit multiperiod post-intervention data typically focus solely on the dynamics of responses to the intervention. If aggregate shocks affect the returns to the intervention, however, it can be difficult to identify true response dynamics. A number of studies have tracked the effectiveness of "Graduation" programs designed to improve the welfare of the very poor in developing countries (Bandiera *et al*., forthcoming; Banerjee *et al*., 2015; Blattman *et al*, 2016)) without careful attention to the evolution of aggregate conditions. Tjernström *et al*. (2013) exploits the timing of the rollout of a program to distinguish the dynamic program effect from aggregate shocks, but identification relies on the assumption that the dynamic effect is additively separable from the effect of the aggregate shocks.

Another common feature of existing studies is that standard errors are computed based on sampling error exclusively. It is well-known in the econometrics literature, however, that the confidence intervals around single-period cross-sectional parameters estimates are too conservative in the presence of time-varying stochastic shocks that are common across cross-sectional units and that alter the parameter estimates themselves (Andrews, 2005). The literature, however, contains few if any estimates of confidence intervals that incorporate the effects of macro variability, as such estimates at a minimum require knowledge of the influence of macro shocks on the estimated treatment effects of interest. These are typically unattainable in a single year study. Thus we lack evidence on the quantitative importance of aggregate shocks for computing confidence intervals and for the assessing the external validity of any single-year estimates.

The issue of the consequences of stochastic variability for gaging external validity may be of particular importance for populations in developing economies, which typically face large exogenous macro shocks. One reason is that a substantial proportion of the labor force in low-income countries is engaged in agriculture, for which weather shocks are very important. This also means that fluctuations in weather in such countries potentially spillover to the non-agricultural sector via product demand and factor price variation. Many developing countries are also single-commodity exporters (e.g., tin, petroleum) and thus are heavily exposed to external price and demand shocks. Populations in these macro environments thus may face very different conditions from year to year.

Recently there have been studies that have examined how a particular estimated relationship varies across different samples of the same population (Allcott, 2015) or across different populations (Dehejia *et al*., 2016; Meager, 2016). These show that generalization of treatment effects across populations can be difficult: in the first case because of the potential for systematic selection bias in the populations that are studied; in the second because of limited information about the underlying heterogeneity across populations and the absence of models that are informative about the factors affecting the behavior under study.

A few recent studies have identified the existence of time-varying returns due to rainfall and seasonality. Attanasio and Augsburg (forthcoming) compare the single-year estimates from an RCT in Anagol *et al.* (2013) to those from additional years to show that returns vary by drought and non-drought years. And Beegle *et al.* (forthcoming) report on findings from an RCT design incorporating interventions in lean and harvest seasons, anticipating that returns will be different in the two seasons. But there is a dearth of studies that have provided estimates of the full inter-temporal distribution of treatment effects or investment returns that are required to inform the decisions of entrepreneurs or policy-makers in a stochastic world. One exception is a recent paper by Zhang (2016), who shows that the estimates of the effects of de-worming in Miguel and Kremer (2004), based on a 38-year distribution of rainfall in the project area, may be as much as twice as high as those that would be obtained in a typical rainfall year, when worm loads would be much lower than those experienced at the time of the experiments.

In the macro literature (e.g., Kroft and Notowidigdo (2016)) it is established that behavioral responses to programs vary over the business cycle, although there is little consensus on the forces that drive business cycles. Hahn *et al.* (2016) distinguish between two types of interactions between the behavior of individuals and aggregate shocks - one in which the shocks are exogenous to agents' behavior and the other in which there is feedback between the choices of agents and the aggregate economic shocks (e.g., changes in unemployment rates). In developing countries, shocks exogenous to individuals, firms and farms are a salient feature of everyday life, and are often readily measurable. Inter-annual and intra-annual (seasonal) variability in weather, variation in international commodity prices, the introduction of new technologies, and even large changes in policies impact decisions and decision outcomes with little or no feedback to the time path of aggregate shocks.

In this paper we obtain estimates of the determinants of the inter-temporal variation in returns to investments in agriculture, in non-farm enterprises, and in human capital from panel data sets from India, Ghana, Sri Lanka and Indonesia to assess the quantitative importance of the influence of exogenous macro shocks, to compare the theoretically appropriate confidence intervals to those based only on cross-sectional sampling variability, and to evaluate the intertemporal external validity of single-year estimates. We focus on investment returns because these are critical parameters for diagnosing the barriers to economic development and thus are essential for policy. For three of our data sets, we are able to quantify a lower bound of the extent of inter-temporal variability in investment returns that is not limited to the time period of the study by identifying important *exogenous* determinants of inter-temporal investment return variability, including intra- and inter-seasonal variability in rainfall and global price shocks.[1]

Small-scale agriculture is the dominant economic activity of the poorest countries. Agricultural production has an important temporal component, typically requiring up front investment of substantial resources in exchange for an uncertain harvest months in the future. There is a substantial literature that describes and models the costly efforts of farmers to reduce this risk or to mitigate its consequences. Despite the centrality of risk to our understanding of the economics of agricultural organization, little attention has been paid to its consequences for the external validity of estimates of the productivity of investments in agriculture. We examine data from Ghana and India in which we can estimate the consequences of variation in aggregate shocks for the external validity of estimates of the return to investment. A primary advantage of starting with agriculture is that we have good *a priori* reasons to believe that the important dimensions of the shocks affecting agrarian returns are measurable.

Based on our estimates of the sensitivity of the investment returns to the realization of rainfall shocks in our agricultural data sets and long time-series of rainfall we also derive lower-bounds on the true confidence intervals of the returns that incorporate the complete distribution of the stochastic shocks and compare them to the standard confidence intervals that are based solely on sampling variability typically reported in single-year studies. We also assess the spatial validity of our estimates of the sensitivity of returns to weather variation and consider whether the realization of macro shocks across spatial units in a single year can be used to identify the effects of intertemporal variations in shocks, exploiting rainfall variation across villages. We also test whether seasonality in rainfall affects the variability in profits for non-farm enterprises, using panel data from an RCT carried out in Sri Lanka, where a large fraction of the labor force is engaged in agriculture. Finally, we quantify the importance of macro price shocks on the returns

[1] Banerjee *et al.* (2014, 2016) provide a general framework for considering the external validity of causal estimates, arguing that researchers should explicitly provide "structured speculation" regarding the scope and degree of external validity of their estimates. Deaton and Cartwright (2016; p.14) focus on RCTs, but acknowledge that the point is general when they argue similarly that causal estimates "must be integrated with other knowledge … if they are to be useable outside the context in which they were constructed." We implement these suggestions and develop an approach that specifies the relevant heterogeneity for generalizing causal estimates in one population at one time to other populations or circumstances, and set down a model of the interactions between those dimensions of heterogeneity and the causal effect under consideration.

to schooling using panel data from Indonesia and use our methods and findings to assess *ex post* the external validity of the returns to schooling in Duflo (2000), which were obtained for a single year.

In section I we present a simple model of the investment decisions of agents in a stochastic environment in which the return to investment is subject to aggregate shocks. The model motivates the IV strategies we and others employ to estimate investment returns and also identifies the information needed to generate decision-relevant confidence intervals.

In section II we present estimates of how rainfall variability affects estimates of the returns to agricultural investments from panel data describing farmers in India and Ghana. We find that in both agricultural settings, the returns to planting stage investments are very sensitive to rainfall realizations and that rainfall realizations are themselves quite variable. Therefore, agricultural profits are highly variable over time in both contexts. Our calculations of lower-bound confidence intervals based on both our parameter estimates of rainfall sensitivity and the actual distribution of rainfall, based on a long time-series of rainfall outcomes, are substantially larger than the standard confidence intervals computed from any single year realization that are based solely on population sampling variability. The lower-bound estimates of the true confidence intervals are in fact quite wide, so that the computed probability that any single year estimate is within a reasonable bound of the expected return facing the farmers is low. We find, for example, that the probability that any single year estimate of the rate of return is within 30 percentage points (on either side) of the expected value of the rate of return is only 5% in both Ghana and India.

Using the Ghana data, where we have rainfall information for 75 villages, we also find that using the cross-sectional variation in rainfall realizations in any one year does not replicate estimates obtained using farmer fixed effects, indicating the necessity of having panel data to identify the influence of rainfall variability on returns, at least in the setting we analyze. We investigate one potential source underlying the discrepancy between the cross-sectional and panel estimates using information on the baseline wealth of the farmers and a 16-year time-series of rainfall. We find that wealth levels are higher in areas with higher mean levels of rainfall, and that in the three panel years after baseline wealth is measured, wealth is significantly correlated with the future weather outcomes in the cross-section. This can generate a positive bias in the Karlan *et al*. (2014) estimates of the consequences of insurance for risk-taking if wealthier households respond more to insurance via a liquidity effect. In the final subsection, we also quantify the extent to which inter-annual output price variation affects estimates of agricultural investment returns. Fluctuations in world commodity prices evidently impact the returns estimates in similar magnitudes to those caused by rainfall variability.

To demonstrate that the sensitivity of returns to investment to aggregate shocks is not limited to the agricultural sector, in Section III we examine the returns to investment in small non-farm service and manufacturing enterprises in Sri Lanka, using the quarterly panel data on micro-enterprises collected by De Mel *et al.* (2008). As we found for farmers, there is significant variation over time in the profits realized by each entrepreneur (cv=0.5), and these fluctuations are consistent with respondents' subjective expectations of profit variability. We estimate a strong seasonal pattern in the returns to investment by these entrepreneurs net of individual illness shocks. This pattern is consistent with respondents' reports that cite aggregate demand fluctuations as an important source of profit variation and are also consistent with the specific seasonality of rainfall in the study area, still highly agricultural, that would alter the intra-annual demand by rural households for non-agricultural products and services.

Although the aggregate shocks driving variation in returns to investment in Sri Lanka are consistent with the seasonality of rainfall in the relevant study areas, we cannot precisely identify the source of demand variation. In section III, we examine the specific macro factors that drive inter-annual fluctuations in short-run measures of the returns to urban schooling in Indonesia, namely fluctuations in the international price of oil and in the rupiah exchange rate. An object of interest in many studies in developed and developing countries is the rate of return to schooling. In the macro growth literature, differences in the rates of return to schooling across countries have been used to quantify the role of schooling in growth (Bils and Klenow, 2000) and to identify skill bias differences in technologies across countries (Caselli and Coleman, 2006). The principal source of data on returns to schooling are from Psacharopolous (1994). This study, however, provides single-year estimates of rates of returns from micro studies by country. To the extent that schooling returns are sensitive to macro shocks, the global variance in these return estimates will overstate the true inter-country variance in expected returns. And estimates of the determinants of cross-country differences in growth or technology choice based on the set of single-year estimates thus may lead to incorrect inferences.[2]

The most influential micro studies recognize the existence of *individual-specific* unobservables that jointly affect schooling and earnings, just as for agricultural and enterprise investments, and attempt to estimate the returns to schooling exploiting policy variation relevant to schooling attainment choices as instruments (e.g., Angrist and Krueger, 1991; Angrist and Krueger, 1992; Card and Lemieux, 2001; Card, 1995; Duflo, 2001). These studies too, however, ignore stochastic variation in earnings returns associated with aggregate shocks. For example, Angrist and Krueger (1991) use variation in the timing of births combined with variation in compulsory schooling laws in the United States as instruments determining schooling and then look at earnings outcomes separately from the 1970 and 1980 Censuses. The differences between the estimated returns to schooling across the two Census years are comparable to or much larger than (in the most complete specification) the difference between the IV and OLS estimates within the same year that is the focus of the study.

There is evidence that aggregate factors have persistent effects on schooling returns. For example, macro shocks, measured by rates of unemployment occurring at the time of labor market entry, evidently affect the returns to schooling, since such shocks differentially affect the initial earnings of college and high-school graduates (Oreopoulos *et al*., 2012). There is also a literature that attempts to identify the sources of longer-term trends in the returns to schooling, but sorting out the causal effects of longer-term economy-wide supply and demand factors has proved difficult in these studies due to the difficulty of identifying the underlying exogenous factors driving fluctuations in the aggregate economy, as is implied by Hahn *et al.* (2016).

We focus on short-term changes in the return to schooling, driven by price and exchange rate fluctuations, since most estimates of schooling returns and studies of their cross-country variation rely on data from a single year. We use panel data on urban wage and salary workers in Indonesia over the period 1993-2008, during which time there were dramatic changes in the

[2] For example, a regression of the rates of return on GDP will be biased to the extent that macro shocks both affect GDP and alter schooling returns (as we show).

world price of oil and a rapid depreciation of the Rupiah as a consequence of the 1998 global financial crisis. We combine data on these shocks with information on industry-specific skill intensity, openness to trade, and the distribution of industries across provinces to estimate the effects of these shocks on the returns to schooling in each year. We find that these external macroeconomic shocks had important effects on the return to schooling as estimated in any year, in line with our expectations of how these shocks affect the relative demand for skilled and unskilled workers. Based on our estimates of the determinants of variation in the returns to schooling and our ability to measure the potentially relevant exogenous macro shocks in Indonesia over a long time period, we revisit Esther Duflo's (2000) study of the effects of the 1973-78 Indonesian school building program and her estimate of the return to schooling. We show that her estimate of the return to schooling based on a single Census year is close to the mean return over the relevant decade, although had she chosen to use another census year the return estimate would have been considerably higher.

In the conclusion, we provide recommendations for the use of theory and external data to improve, or at a minimum assess, the temporal external validity of planned or existing causal estimates.

I. Model

We consider a population of agents indexed by i who each year $t\ (= 1, \dots T)$ make investments whose outcomes are realized subsequently and are uncertain. For simplicty, we divide up each year into an investment period 0 and a realization period 1. Agents invest an amount (a_{i0}) in the first period and realize proceeds of the investment (y_{i1}) in the subsequent period. For agriculture, the first period would correspond to planting, and the second the harvest period. For business owners, the first period may be capital investments or inventory investments and the second the realized demand for the firm's products. For single, one-time investments like schooling, the second period would consist of the post-schooling career with the outcomes (earnings), but not investments, varying year to year, with corresponding changes in the notation.

We focus our initial attention on the repeated investments of agents who operate farms and firms. In any year, the observable second-period profit of the of the agent is given by the function

$$\pi(a_{i0}, s_1, \lambda_i, \epsilon_i) = \epsilon_i y_1(a_{i0}, s_1) + \lambda_i - a_{i0}. \tag{1}$$

Profits are determined by first-period investments, an aggregate shock s_1, an agent fixed effect λ_i that reflects unobserved variation in the level of output (harvests, sales) across agents, and an unobserved idiosyncratic agent or firm characteristic ϵ_i that we assume has distribution $g(\epsilon)$ independent of both s_1 and observed agent characteristics, but which is known to the agent. We omit other observed characteristics of the agent or firm for notational convenience. The standard evaluation problem is that, of course, we can never simultaneously observe different levels of investment for the same agent *i.* That is not our focus in this paper, so we will assume that we have (a) panel data with $T \geq 2$; and (b) an appropriate instrument (ζ_{it}, as discussed below) to disentangle any endogeneity of first-period investment (a_{i0}) due to the idiosyncratic variation ϵ_i, which is unobserved by the analyst.

The problematic source of uncertainty for the agent in the model is the random vector s_1, the state of nature in the second period, which has distribution $f(s_1|\Omega_{i0})$, where Ω_{i0} is the agent's information set at the time of the choice of a_{i0}. While *s* may be multidimensional, we will initially restrict our attention to a single dimension. For agriculture, this could be the total rainfall received during the growing period; for firms the number of product buyers or the output price.

The agent chooses her investment to solve

$$\max_{a_{i0} \in A} E_{s_1}[V_{i1}(\pi(a_{i0}, s_1, \lambda_i, \epsilon_i), z_i)|\Omega_{i0}] \tag{2}$$

V_{i1} is an increasing and concave function of π_i and $z_i \in \zeta_i$ is a characteristic of *i* that might affect that agent's valuation of a realization of profits in a particular state. If the agent is risk neutral, then decisions depend only on the expected value of the return to first-period investment

$$0 = \int \epsilon_i \frac{\partial y_1(a_{i0}, s_1)}{\partial a_{i0}} f(s_1|\Omega_{i0}) ds_1 - 1 \tag{3}$$

More generally, however, a risk-averse or credit-constrained agent will base decisions on the expected marginal utility of profits,

$$0 = \int \frac{\partial V_{i1}(\pi, z_i)}{\partial \pi} \left(\epsilon_i \frac{\partial y_1(a_{i0}, s_1)}{\partial a_{i0}} - 1 \right) f(s_1|\Omega_{i0}) ds_1 \tag{4}$$

and therefore on the complete distribution of profits over aggregate shocks. If $\frac{\partial^2 V_{i1}(\pi,z_i)}{\partial\pi\partial z_i} \neq 0$, then by the implicit function theorem we can define the function $a_{i0}(\zeta_i, \epsilon_i)$ describing i's choice of the first-period level investment to satisfy (2).

To an analyst attempting to estimate the profit function, variation in investment choice by otherwise similar agents is required in order to measure the return to investment. The estimation problem is that optimal investment of any agent *i* will be related to ϵ_i, which is not observed by the analyst. z_i is a candidate for an instrumental variable, or for an intervention that induces a change in a_{i0} if z_i is uncorrelated with all unobserved individual agent characteristics. This will be the case, for example, where z_i is a vector of randomized assignments to treatment groups receiving varying grants of cash, as is used in two of our examples below.

Similarly, if $\omega_{i0} \in \zeta_i$ represents an informative signal regarding $f(s_1|\Omega_{i0})$ that varies across areas, then ω_{i0} is another candidate for an instrumental variable that generates variation in a_{i0} if the signal is independent of unobserved agent characteristics. This will be the case in the first example below, where ζ_i contains an informative forecast of seasonal rainfall.

The marginal return to investment for agent i in aggregate state s_1 at investment a_{i0} is

$$\beta_i(a_{i0}, \epsilon_i, s_1) \equiv \frac{\partial\pi(a_{i0}, s_1, \lambda_i, \epsilon_i)}{\partial a_{i0}} = \epsilon_i \frac{\partial y_1(a_{i0}, s_1)}{\partial a_{i0}} - 1. \tag{5}$$

We typically have data on $\{a_{i0}, y_{i1}, \zeta_i, s_1\}_{i\in I}$. Suppose that there is variation in ζ_i sufficient to generate variation in $\{a_{i0}(\zeta_i, \epsilon_i)\}$ across the observed agents such that we observe investments over the full range $a_{i0} \in A$ and a sufficiently large cross sectional sample that for each value $\hat{\zeta}_i$ we observe

$$\begin{aligned} \hat{\beta}\left(\hat{a}_0(\widehat{\zeta_i}), \hat{s}_1\right) \approx \beta(\hat{a}_0, \hat{s}_1) &= E_{\epsilon_i}\beta_i\left(a_{i0}\left(\hat{\zeta}_i, \epsilon_i\right), \epsilon_i, \widehat{s_1}\right) \\ &\equiv \int \epsilon \frac{\partial y_{i1}\left(a_{i0}(\hat{\zeta}_i, \epsilon), s_1\right)}{\partial a_{i0}} g(\epsilon) d\epsilon - 1 \end{aligned} \tag{6}$$

for $s_1 = \hat{s}_1$, the aggregate shock realized in our data.

Even in this benign scenario, in which we can identify the expected return (over agents) to investment $\beta(\hat{a}_0, \hat{s}_1)$ for any level of investment, it remains the case that we observe this return function at only a single value of the aggregate shock s_1. This information may be of relatively limited value because it is typically the entire distribution of returns across possible aggregate states that is relevant for policy decisions or for agent investment choice.

Two aspects of the random vector s_1 are particularly important for assessing the intertemporal external validity of single-year estimates and the bias in standard errors based solely on population sampling variability. First, the variability in s_1. A single estimate of the average rate of return to an action is more informative in an environment in which the variability of s_1 is more limited. Second, the responsiveness to realizations of s_1 of the marginal value of increasing a_{i0}. For given variability of s_1, the larger is this responsiveness, the less informative a single estimate of the average rate of return.

One approach to quantifying the importance of aggregate shocks for external validity is to calculate the probability that any particular point estimate of the returns to investment is within a pre-specified interval around the expected value of that return. Consider a unidimensional continuous shock. We observe $\beta(a_0, s_1)$, asymptotically without error, for $a_0 \in A_0$, but only for $s_1 = \hat{s}_1$. For any particular value of $a_0 = \hat{a}_0$, how likely is it that our estimate of the return to a_0 is within δ of $E_{s_1}\beta(\hat{a}_0, s_1)$? To simplify notation, define

$$\frac{\partial^2 \pi(\hat{a}_0, s_1)}{\partial a_0 \partial s_1} = \frac{\partial \beta(\hat{a}_0, s_1)}{\partial s_1} = \alpha(a_0, s_1) \tag{7}$$

If $\alpha(a_0, s_1)$ is a constant (so $\beta(a_0, s_1)$ is linear in s_1), then we write it as α and

$$\begin{aligned} &prob\left[\left(E\left(\beta(\hat{a}_0, s_1)\right) - \delta\right) \leq \beta(\hat{a}_0, s_1) \leq \left(E\left(\beta(\hat{a}_0, s_1)\right) + \delta\right)\right] \\ &\qquad = prob\left[\left(E\left(\beta(\hat{a}_0, s_1)\right) - \delta\right) \leq E\left(\beta(\hat{a}_0, s_1)\right) + \alpha\left(s_1 - E(s_1)\right) \right. \\ &\qquad \left. \leq \left(E\left(\beta(\hat{a}_0, s_1)\right) + \delta\right)\right]. \end{aligned}$$

Therefore,

$$\begin{aligned} &prob\left[\left(E\left(\beta(\hat{a}_0, s_1)\right) - \delta\right) \leq \beta(\hat{a}_0, s_1) \leq \left(E\left(\beta(\hat{a}_0, s_1)\right) + \delta\right)\right] \\ &\qquad = prob\left[-\frac{\delta}{\alpha} \leq \left(s_1 - E(s_1)\right) \leq \frac{\delta}{\alpha}\right] \end{aligned} \tag{8}$$

For any given value of δ, a larger standard deviation of *s* or a larger cross derivative $\frac{\partial^2 Y_1(\hat{a}_0, s_1)}{\partial a_0 \partial s_1}$ reduces the probability that a given observation of the return to an action is near the expected value of that return. If s is normally distributed, then the probability of estimating a return to *a* that is within $\alpha\sigma_s$ of its expected value is approximately 68%.

There are multiple dimensions of aggregate uncertainty that potentially influence the returns to investment. For k-dimensional $\boldsymbol{s}$, (8) generalizes with $\boldsymbol{\alpha}(\hat{a}_0, \boldsymbol{s}_1) \equiv D_{s_1}\beta(a_0, \boldsymbol{s_1})$ to

$$\begin{aligned} &prob\big[\big(E\big(\beta(\hat{a}_0, \boldsymbol{s}_1)\big) - \delta\big) \leq \beta(\hat{a}_0, \boldsymbol{s}_1) \leq \big(E\big(\beta(\hat{a}_0, \boldsymbol{s}_1)\big) + \delta\big)\big] \\ &\quad = prob\big[-\delta \leq \boldsymbol{\alpha}'\big(\boldsymbol{s}_1 - E(\boldsymbol{s}_1)\big) \leq \delta\big] \end{aligned} \tag{9}$$

What if, as is likely, some dimensions of the aggregate state vector s are unobserved or are ignored? Then it is straightforward to show that the confidence interval estimated based on the variation in any subset of s always *understates* the confidence interval that would be calculated incorporating all k dimensions. This is true regardless of the sign of the correlation between the omitted and included macro shocks. We illustrate this point with a two-dimensional example. Normalize the two-dimensional state vector $\boldsymbol{s}_1 = (s_a, s_b)'$ so that the measure s_i of each dimension's realization has mean 0 and variance 1. The correlation between the two is ρ. Equation (9) describes the probability that the estimate of the return to investment in a given realization of the aggregate state lies within δ of the expected value over states of the return. The variance of $\boldsymbol{\alpha}'\big(\boldsymbol{s}_1 - E(\boldsymbol{s}_1)\big) = \boldsymbol{\alpha}'\boldsymbol{s}_1$ is $\alpha_a^2 + \alpha_b^2 + 2\alpha_a\alpha_b\rho = \sigma_T^2$.

Suppose, however, that we do not observe s_b. We therefore estimate

$$\tilde{\alpha}_a = \alpha_a + \alpha_b\rho \tag{10}$$

and calculate

$$prob[-\delta \leq \tilde{\alpha}_a s_a \leq \delta] = prob[-\delta \leq (\alpha_a + \alpha_b\rho)s_a \leq \delta]. \tag{11}$$

While $\tilde{\alpha}_a s_a$ has the (correct) mean of 0, it's variance is only $\alpha_a^2 + \alpha_b^2\rho^2 + 2\alpha_a\alpha_b\rho = \sigma_M^2$. The difference between the true and estimated variance in this case is

$$\sigma_T^2 - \sigma_M^2 = \alpha_b^2(1 - \rho^2). \tag{12}$$

Thus, the omission of the unobserved dimension of aggregate variation always causes us to overestimate the probability that any estimate of β, given a particular realization of the aggregate shocks, is near the expected value of β. Our corrections of confidence intervals around estimates of the returns to investment to account for some dimensions of aggregate variation, therefore, are always lower bounds to the true confidence intervals if there remain any unobserved sources of aggregate variation in the returns to investment that are not perfectly correlated with the dimensions we observe.

We need not restrict our attention to the expected value of the return to investment across aggregate states. Provided we have a credible estimate of α in the linear case, or more ambitiously of $\alpha(s)$ in the more general case, these estimates can be combined with estimates of the distribution of states to provide an estimate of the full distribution of returns to investment. For agriculture, a key risk element is rainfall, for which in many countries there are multiple observations over time and space. For this sector therefore, given the stationarity of rainfall realizations, we can plausibly characterize the full distribution of returns due to this source of stochastic variation that face farmers and to calculate lower-bounds on confidence intervals that take into account stochastic variation in investment outcomes. In particular, we observe the realization $\hat{\boldsymbol{s}}_1$ of $\boldsymbol{s_1}$ (total rainfall and an index of its distribution over the season for Ghana; and total rainfall in India) over many years, and thus have an estimate $\hat{f}(\boldsymbol{s_1})$.[3] We use repeated draws of $\hat{\boldsymbol{s}}_1$ from $\hat{f}(\boldsymbol{s_1})$ and $\widehat{\boldsymbol{\alpha}}$ to simulate the distribution of returns to investment, $\widehat{\boldsymbol{\alpha}} \cdot \widehat{\boldsymbol{s_1}}$. Because we are observing only a subset of the dimensions of aggregate uncertainty, as noted the simulated distribution of returns to investment will provide a lower bound to the true variance of returns.

Finally, another source of aggregate variation in investment returns is price variability. Appendix A discusses output prices. The effect of variation in output prices on the returns to investment has a simple, linear form that we use to discuss its effect on the temporal external validity of estimates of the return to investment in agriculture in India and Ghana.

II. RETURNS TO AGRICULTURAL INVESTMENT

A. The Return to investment in Indian ICRISAT Villages 2005-2011

We begin by using seven years of panel data on farmers from the ICRISAT Village Dynamics in South Asia (VDSA) surveys for the years 2005-2011 to estimate the returns to planting-stage investments, their sensitivity to rainfall realizations, and the corrected shock-inclusive confidence intervals. Here we use the insight that exogenous changes in expectations can be used to identify investment returns and exploit information on the official annual

[3] The distribution of total rainfall in the semi-arid tropics of India and West Africa is well characterized as iid over years (e.g., Manzanas *et al*. 2015))

forecasts of monsoon rainfall issued by the India Meteorological Department (IMD). Suppose that ω_{i0} contains information about s_1; for example, that increases in ω_{i0} are associated with a first order stochastic dominant shift in the distribution of s_1 and that s_1 and a_{i0} are complements in production (as for example in a model with multiplicative shocks). A risk-neutral farmer increases investment upon receipt of a forecast of good weather:

$$\frac{da_{i0}}{d\omega_{i0}} = -\frac{\frac{\partial E_{s_1}\left(\beta(a_0, \zeta_i, s_1) \middle| \Omega_{i0}(\omega_{i0})\right)}{\partial \omega_{i0}}}{\frac{\partial E_{s_1}\left(\beta(a_0, \zeta_i, s_1) \middle| \Omega_{i0}(\omega_{i0})\right)}{\partial a_{i0}}}. \tag{13}$$

The denominator is negative, so $sign\left(\frac{da_{i0}}{d\omega_{i0}}\right) = sign\left(\frac{\partial \beta_i}{\partial \omega_{io}}\right)$. As we would expect, an informative signal that the realization of the aggregate shock is likely to be larger leads to greater (less) investment if the shock and investment are complements (substitutes).

The semi-arid tropics where the ICRISAT villages are located are one of the most difficult environments to farm, as rainfall is both relatively low on average and highly variable. The data we use are based on surveys of farmers in each of the six villages from the first generation ICRISAT VLS (1975-1984) but over the years 2005-2011. The villages are in the states of Maharashtra (four) and Andhra Pradesh (two). The ICRISAT data set contains farmer-level investment and profit data for seven consecutive years, permitting us to quantify the sensitivity of investment returns to aggregate shocks.

There are three additional features of the ICRISAT data relevant to our investigation. First, input and output information is provided in approximately three-week intervals collected by resident investigators. This enables us to precisely measure investments made within a season prior to the realization of rainfall (s_1 in the model) as well as the season-specific profits associated with those investments. Second, there are data on daily rainfall for each of the six villages for as long as 26 years. This enables us to both estimate the influence of rainfall realizations on investment returns and to characterize the distribution of rainfall states *f*(*s*) faced by farmers so that we can compute confidence intervals that take into account stochastic outcomes as well as population sampling variability. Rainfall and thus profit variability in the ICRISAT villages is high.[4]

[4] Our measure of profits is the value of agricultural output minus the value of all agricultural inputs, including the value of family labor and other owned input services.

The average coefficient of variation in profits over the period experienced by the ICRISAT farmers is 1.4.

The third feature of the ICRISAT data that we exploit is that for the four villages in Maharashtra, the forecasts of monsoon rainfall issued by the India Meteorological Department (IMD) in late June have been moderately successful in predicting *kharif*-season (July-September) rainfall and, most importantly, significantly influence the planting-stage decisions of the ICRISAT farmers (Rosenzweig and Udry, 2014) in the July-August planting stage. We use the IMD forecast for the southern peninsula as an instrument in an IV strategy for estimating the net returns to planting-stage investments (the value of labor used in plowing, seeding and fertilizing plus the costs of the material inputs) and their sensitivity to rainfall variation when such investments are endogenous. In the context of the model, the IMD forecast is a signal to farmers (ω_{i0}) that affects their expectations of seasonal rainfall and thus profitability but, for given investments, rainfall realization and prices, should have no direct effect on profitability. We estimate a conditional profit function using the ICRISAT panel data, treating planting-stage investments as an endogenous choice that responds to the rainfall forecast.

To identify investment returns, we first have to consider how a forecast that is common to all farmers in a given year affects investment decisions differentially across farmers. We take advantage of the fact that the ICRISAT data set has detailed information on soil properties – soli color, depth, etc. - that not only affect agricultural profits directly but also mediate how rainfall variation affects profits. Therefore, we follow the literature (e.g., Carter and Lyybert, 2012) and permit interactions between rainfall and soil characteristics to affect profits. If soil characteristics affect how rainfall impacts on profitability, then farmers' investment decisions will also respond differentially to a rainfall forecast depending on their soil characteristics.

We impose additional structure to ensure that the forecast instruments satisfy the exclusion restriction. Agricultural profits depend on investments in planting-stage inputs and on the realization of rainfall, and as our model has emphasized, on the interaction between these. There is also good evidence (Sharma and Acharya 2000) that profits depend as well on lagged rainfall (differentially depending upon farm characteristics, particularly soil depth) through the soil moisture overhang effect. Hence we specify a linearized version of the farm profits of household h in village v in year t that is quadratic in planting-stage investments and includes interactions between current-season rainfall and rainfall in the prior year with a set of farm-specific land characteristics, a farmer fixed effect, and village-year fixed effects that absorb time-varying village-specific input prices (particularly wages) that could be correlated with rainfall forecasts. A key feature of our specification is that it also allows the effects of planting-stage investments on profits to depend on the realization of rainfall. Excluded from the profit specification are the rainfall forecast and its interactions with exogenous fixed land characteristics. [5]

Table 1 reports fixed-effects instrumental variable (FE-IV) estimates of the profit function, with the FE at the farmer and village-year levels. The IMD forecast interacted with the characteristics of the farm and farmer are the instruments for planting-stage investments. All profit function specifications include the rainfall variables interacted with total landholdings, irrigated landholdings, soil depth, and four soil types (red, black, sandy, loam). The first column of Table 1 reports estimates from the profit specification that is linear and quadratic in investment. Based on those estimates we can strongly reject the hypothesis that larger planting-stage investments do not increase profits over almost the full range of the investment distribution in the sample. In the second column we add rainfall interactions with the investment variables and the interactions of investments with rainfall and rainfall with four soil types (two indicator variables for depth, loam and sandy (left out are black and red soil). We can also reject that investment returns do not depend on rainfall (p=.031) and that the effects of rainfall on investment returns do not depend on soil characteristics (p=.032). These estimates thus imply that *ex-post* optimal investments depend on realized rainfall outcomes, or, put differently, how much under-investment one would infer from profit function estimates depends on what is assumed to be the typical rainfall outcome. The estimates in column 2 imply that at mean levels of investment in the ICRISAT sample, returns to planting-stage investment are positive over the full range of rainfall realizations observed in the data, as shown in Figure 1.

[5] This omission is the primary identification assumption required to estimate the returns to planting-stage investments. That is, conditional on realized rainfall (or village-year fixed effects) the forecast of total rainfall in the monsoon affects profits only through its effect on investments. Additional details are provided in Appendix B.

The point-wise confidence intervals depicted in Figure 1 for the rainfall-specific returns to agricultural investment assume that the only source of variation in the estimates is sampling error. However, our estimates suggest that variability in rainfall also strongly affects returns. Our primary goal is to understand the degree to which an estimate of the return to agricultural investment in a particular season, conditional on a particular aggregate weather realization, provides information about the expected value of this return. We do this first assuming that we know α exactly, using our estimate of alpha from column 2 of Table 1 so that we observe the rate of return precisely given any specific weather realization. Thus the confidence interval for any estimate obtained for a given rainfall realization s_{1t} will reflect only the variability in the distribution of s_1, that is, $f(s_1)$. To construct the confidence interval based solely on the exogenous variation in rainfall, and to assess the probability that any one estimate is within pre-specified interval around expected profitability (profit at mean rainfall) we use the rainfall time series from Kinkheda, the ICRISAT village with the longest continuous history - 26 years. Standard tests of normality indicate non-rejection of the null (e.g., Shapiro-Wilk W test [p=.28]). We thus use the standard deviation and mean of the rainfall distribution to assess the confidence interval inclusive of rainfall variability. Based on the actual rainfall distribution parameters and α, we draw 1000 profit returns estimates. The distribution of profit returns, taking into account rainfall variability, is shown by the line in Figure 2.

The solid line characterizing the distribution of profit returns does not take into account sampling variability, that α is an estimate. We thus computed the distribution of returns combining the standard deviation of our estimates (due to sampling variability) and rainfall variability, assuming that the two errors are independent. That distribution is depicted by the cross-hatched line in Figure 2. We contrast these with the distribution of estimates implied by sampling variability alone, as reported in all studies, for two returns estimates, at the 25th and 75th percentile of the rainfall distribution. As can be seen, the confidence intervals based solely on sampling variation again severely understate the confidence interval constructed from the distribution of estimates incorporating both sampling variation and stochastic variability in rainfall. And, as we have shown, our rainfall-inclusive confidence-interval estimate is itself a lower bound of the true confidence interval incorporating all of the macro shocks facing farmers.

Based on the estimated distribution of returns, we can compute, based on (3), the probabilities that any one estimate of returns obtained in a random, single year lies within some interval around the true expected (over the rainfall distribution) investment return. The first column of Table 2 reports these probabilities for the Indian village, for intervals ranging from 10 percentage points to 50 percentage points on either side of the mean. Note that these probabilities represent upper bounds for two reasons. First, to emphasize the role of aggregate shocks we have assumed that there is no sampling error in this table, unlike for Figure 2. Second, these probabilities are solely based on rainfall variability; we have shown that the incorporation of any other aggregate shocks would reduce these probabilities further.[6]

B. The Return to investment in Northern Ghana, 2009-2012

We use three years of panel data from 1,352 households in 75 communities in the Northern Region of Ghana to examine the variability in the returns to planting stage investments as a function of weather realizations over these three years across this region. The large number of communities and their broad spatial extent raises the possibility of using cross-sectional variation in weather realizations to estimate α; we can thus use these data to compare cross-sectional and fixed effect estimates of the responsiveness of returns to weather shocks. This data set contains detailed plot level information on inputs at each stage of production and seasonal output, along with geographical information on the location of each plot that can be combined with data on weather realizations over the season. An important advantage of these data is that they were collected in the context of a randomized controlled trial that varied the availability of rainfall index insurance and substantial grants of cash across the sample of farmers and over time.[7]

Randomized grants of insurance or cash (z_{i0}) affect the choice of a through their effect on $V_{i1}(\quad)$. In this experiment, grants of cash or (especially) insurance reduced marginal utility in states with low returns to investment relative to states with high returns on investment, so

[6] To the extent that rainfall variability is correlated with other shocks that impinge on profits (e.g. pests, temperature, price), we cannot say that we have identified that part of returns variation due solely to rainfall. That fact is immaterial for the current investigation, but it affects the interpretation of the $\hat{\alpha}$ coefficients, which could be salient for other purposes.

[7] See Karlan *et al*. (2014), section III and online appendix 1 for a detailed description of the sample, data collection procedures, index insurance and cash grants interventions and the randomization.

$$\frac{da_{0i}}{dz_{i0}} = -\frac{\int \frac{\partial^2 V_{i1}(\pi, z_{i0})}{\partial \pi \partial z_0} \frac{\partial \pi}{\partial a_{i0}} f(s_1|\Omega_{i0}) ds_1}{\int \left[\frac{\partial^2 V_{i1}(\pi, z_{i0})}{\partial \pi^2} \frac{\partial \pi}{\partial a_{i0}} + \frac{\partial V_{i1}(\pi, z_{i0})}{\partial \pi} \frac{\partial^2 \pi}{\partial a_{i0}^2}\right] f(s_1|\Omega_{i0}) ds_1} \geq 0 \qquad (14)$$

for risk-averse farmers.

Karlan *et al*. (2014) found that the randomized availability of rainfall index insurance to farmers and the randomized receipt of a cash grant in the sample generated a strong investment response.[8] The randomized allocation of this sample into a variety of subsamples facing exogenously varying budget constraints provides us with exogenous variation in planting season investments which we use to estimate the return to these investments across varying realizations of weather.

There is a great deal of variability in rainfall and thus profits in northern Ghana, as in the ICRISAT sample of farmers. The average standard deviation of profits (the value of harvest minus the cost of all purchased inputs and the value of family labor) for each farmer over the three years is \$450, while mean profits over the three years is \$121.[9] We create a two-dimensional indicator of the weather shock in each community in each year $\boldsymbol{s_{1vt}} = (R_{vt}, I_{vt})$ using information on the total rainfall (in millimeters) received that year and by the index designed for rainfall insurance. The insurance product provides an index (I_{vt}) of the weather realization that combines information on the amounts and timing of daily rainfall to predict harvest for the most important single crop in the region, which is maize. The index is constructed from daily rainfall data available over the period 1983 – 2013 that we obtained from the Ghana Agricultural

[8] Farmers who were insured with rainfall index insurance increased their investment in cultivation by an average of \$266 (s.e. \$134) over a baseline expenditure of approximately \$2058. Expenditure on fertilizer increased by \$56 (s.e. \$17) upon receipt before the cultivation season began of a cash grant averaging \$420.

[9] The fact that Karlan *et al*. find strong responses of investment to access to rainfall index insurance provides *a priori* evidence that investment returns vary by rainfall. They find in addition that the treatment effect in a given year of having access to insurance on harvest values is higher for households in areas that receive higher rainfall in that year. Karlan *et al*., however, do not directly investigate the effect of weather realizations on the rate of return to investment, nor do they explore the implications of the sensitivity of investment returns to weather for the distribution of expected returns.

Insurance Pool (who have developed and market the successor rainfall index insurance product).[10]

Plot profits depend on planting season investments, on the realization of s_{1vt}, and on their interaction. The planting season investments we examine include clearing, field preparation and fertilizer application. We use the random assignment of households to varying treatments of cash grants and grants of or subsidies to rainfall index insurance, and interactions of these treatments with baseline plot area and land characteristics as instruments for plot level investment. Appendix B contains additional discussion of the estimation procedure.

We begin by characterizing the variation in the returns to planting season investment over both time and across space in northern Ghana. We divide the 75 communities in the sample into 11 geographic clusters based on their proximity to a TRMM grid point and estimate the returns to planting season investment separately for each cluster-year, with no effort to associate any variation in the return to investment with weather realizations in that cluster-year: for farmer *i* in cluster *v* in year *t*,

$$\pi_{ivt}=\beta_{vt}a_{ivt} + \lambda_{iv} + \epsilon_{ivt}. \tag{15}$$

The results $(\hat{\beta}_{vt})$are reported in Figure 3 as diamonds. The community clusters are identified by the integers on the horizontal axis, and there is a separate estimate of the return to investment for each year and each cluster. The black lines show 95% confidence intervals around each estimate based solely on sampling variation. The return to investment is 34%, averaged across cluster years. However, it is immediately apparent that the estimated return to planting season investment varies dramatically both across and within clusters over time, and we can resoundingly reject the hypothesis that the returns are equal across all 32 cluster-years ($\chi^2(31)$ = 101.1, p=0.000).

Because these are (unbiased) estimates of the returns to investment in their individual cluster-years, the variance across them overstates the true variance in mean investment returns

[10] We verify the accuracy of these data for the second half of this period using the Tropical Rainfall Measurement Mission at a 0.25 x 0 .25 degree resolution "The TRMM and Other Data Precipitation Data Set, TRMM 3B42," Huffman and Bolvin (2015).

across cluster-years (James and Stein (1961)).[11] The variance across cluster-years of our estimates of the return to investment is 4.71 (implying a coefficient of variation of 609%), but the average sampling variation of our estimates is 2.48. An unbiased estimate of the true variance in mean investment returns across cluster-years, then, is 2.24 (coefficient of variation =420%).

Accordingly, we implement a pair of empirical Bayesian shrinkage estimators to move the individual estimates of cluster-year returns towards their grand mean, reducing $E\left(\hat{\beta}_{vt}-\beta_0\right)^2$ at the cost of biasing the estimates of returns. The simple James-Stein shrinkage estimates (Efron (2012)) appear as green circles in Figure 3.[12] The Hudson-Berger extension of the shrinkage estimator to account for unequal variances of the estimated returns (Morris and Lysy (2012)) are reported as red triangles.[13] These shrinkage estimators provide conservative estimates of the degree of variation in returns across cluster-years – the variances of these estimates of the returns are, respectively, 1.63 and 1.14 – but the variation remains large (coefficients of variation of 373% and 311%).

To take a typical example, in community cluster 2 the unbiased estimates of the return to planting season investment over the three years are 54%, -192%, and +248%. The James-Stein estimates of the same sequence are 32%, -112% and +146%; the Hudson-Berger estimates are 51%, -166% and 229%. Importantly, the magnitude of the variation in returns to planting season investment over time within a cluster appears to be larger than the variation over clusters within a year, for each of the three estimators. The mean across clusters of the standard deviation of

---

[11] Let v denote a cluster of villages and t identify the year. Define the cluster-year return to investment as $\beta_{vt} \equiv \frac{1}{N_{vt}}\sum_i \beta(\alpha_i, s_{vt}) = \beta_0 + z_{vt}$ where i indexes the N_{vt} farmers cultivating in cluster v at year t. α_i is investment by farmer i and s_{vt} is the rainfall realization in cluster v during year t. β_0 is the mean return to investment in the population. The mean zero random variable z_{vt} summarizes the variation across cluster-years in average returns to investment. The estimates plotted in Figure 3 are $\hat{\beta}_{vt} = \beta_0 + z_{vt} + v_{vt}$ where v_{vt} is sampling error. The variance in investment returns across cluster years is $E(z_{vt}^2) = E\left(\hat{\beta}_{vt} - E\left(\hat{\beta}_{vt}\right)\right)^2 - E(v_{vt}^2)$, where the expectation is taken across cluster-years. Thus the observed variation in returns to investment across cluster-years overstates the underlying variance of returns.

[12] $\hat{\beta}_{vt}^{JS} = \bar{\beta} + \left(1-\frac{c}{\hat{F}}\right)\left(\hat{\beta}_{vt}-\bar{\beta}\right)$ with $\bar{\beta} = \frac{\sum\hat{\beta}_{vt}}{VT}$, $\hat{F} = \frac{(N-VT)\sum\left(\hat{\beta}_{vt}-\bar{\beta}\right)^2}{SVT}$ and c is the value of $F(N, VT-N)$ that sets a significance level of 0.05, N is the sample size, VT is the number of cluster-year returns estimated, and S is the total sum of squared errors.

[13] $\hat{\beta}_{vt}^{JS} = \bar{\beta} + \left(1-\hat{\lambda}_{vt}\right)\left(\hat{\beta}_{vt}-\bar{\beta}\right)$ where $\hat{\lambda}_{vt} = \frac{(VT-2)\hat{\sigma}_{vt}^2}{\sum_{vt}\left(\frac{\hat{\beta}_{vt}}{\hat{\sigma}_{vt}^2}\right)^2}$. Where $\hat{\sigma}_{vt}^2$ is the sampling variance of $\hat{\beta}_{vt}$.

returns over time within each cluster is 1.21, 1.14 or 0.93 while the standard deviation across clusters of the mean (over time) of returns within each cluster is 0.58, 0.21 or 0.35 and (unbiased, James-Stein, Hudson-Berger estimates, respectively).

What drives this dramatic variation in estimated returns to investment across these cluster-years? We examine the influence of our two measured dimensions of weather realizations on realized returns to planting season investments in Table 3. Table 3 reports the results from household fixed effects instrumental variable estimates of the profit function using all three years of the panel and the randomized treatments as instruments. The first column estimates the average effect of planting-stage investments on profits, and shows that the average return to planting-stage investment is approximately 100%. In column 2, we report estimates from a specification that permits interactions between planting-stage investments and the vector of weather realizations; these are our estimates of $\boldsymbol{\alpha}$. These interactions are jointly significantly different from zero ($\chi^2(2) = 6.32\ p = 0.04$) and imply that the returns to planting-stage investments at the sample mean weather realization are imprecisely measured but quite high ($r =$ 0.96 s.e. 0.72), and that returns are strongly positively affected by improvements in the timing ($r = 0.12$, s.e.$= .06$) and amount ($r = 0.03$; s.e.$= 0.01$) of rainfall during the growing season.

Figures 4A and 4B display the returns to planting stage investments over the range of each of the dimensions of rainfall realizations in the data set based on the estimates in column 2 of Table 3. The pointwise confidence intervals presented in Figure 4 are again based on the assumption that sampling error is the only source of uncertainty in the estimate of the return to investment. However, our estimates from the Ghana data, as for the India data, imply a very strong response of net returns to the realization of weather in any season. In order to investigate the implications of weather variability in the Ghana context for the computation of confidence intervals, we use the 65 years of rainfall data that are available from the historical records of the Northern Regional weather station in Tamale to estimate the joint distribution of (R_{vt}, I_{vt}).[14] We

[14] We cannot reject normality of the distribution of R_{vt} for any of the communities in the TRMM data. There is no evidence of serial correlation in total rainfall in the savanna zone of Ghana (Manzanas *et al.*, 2014), so we parameterize total rainfall in any community as a draw from a normal distribution with a mean and standard deviation equal to our the mean and standard deviation of total rainfall over the growing season recorded at Tamale between 1944 and 2008. The weather index I_{vt} takes on nine values, depending upon the distribution of rainfall

then use draws from this estimated distribution to examine the variability in the realized returns to investment that will be generated by the variability in weather conditions in northern Ghana.

We begin by again assuming that the variation in weather across communities and over the three years of the Ghana sample has provided us with sufficient information to estimate the dependence of these rates of return on weather realizations (the vector $\boldsymbol{\alpha}$) precisely. Based on our estimates of the rainfall distribution and $\boldsymbol{\alpha}$, we can simulate the distribution of realized returns to investment. This distribution is reported as the solid line in Figure 5. Of course, we only estimate $\widehat{\boldsymbol{\alpha}}$, so Figure 5 also reports the distribution of expected returns taking into account both rainfall variability *and* the sampling error in our estimates $\widehat{\boldsymbol{\alpha}}$**.**

In contrast, Figure 5 also reports as we did for the Indian sample the distribution of expected returns generated by sampling error alone at two specific realizations of $s_1 = (R, I)$ – "good rain" that generates an expected net return equal to the 75th percentile of the overall distribution of net returns, and "poor rain" that generates an expected net return equal to the 25th percentile of the overall distribution. As can be seen, consistent with our findings from the Indian data, the confidence interval one would calculate from the standard errors of the net income function at either of these two specific realizations would dramatically understate the width of the confidence interval constructed from the distribution of estimates that incorporates both sampling variation and the underlying variability in weather conditions.

Given our estimates of the distribution of weather realizations and $\widehat{\boldsymbol{\alpha}}$, we can use (9) to calculate a lower bound of the probability that any single estimate of the returns obtained in a random season drawn from this rainfall distribution lies within some distance δ of the true $ex\ ante$ expected return to investment. These probabilities are reported in the second column of Table 2, for $\delta \in (10, 20, \dots 50)$ percentage points. Interestingly, despite the very different settings that the Ghana and India ICRISAT farmers inhabit and the different IV strategies used to obtain the α's, these lower-bound confidence intervals are almost identical for each sample. In both samples, for example, the estimates indicate that the probability that a single-year's estimate of profit returns lies within 10 percentage points on either side of the expected return

across days in the season. Again we calculate this index for each year between 1944 and 2008 and we estimate the probability of realization of each of these values separately for each of the four quartiles of the overall rain distribution.

is just a little more than 1 percent for the Indian village farmers and 2 percent for the Ghana farmers; and there is only a 6% chance of even being within 50 percentage points on either side of the expected return for any single-year study in both samples. As we have noted, these are upper bounds on the probabilities because we are considering only variation in rainfall. Variation in other dimensions of the aggregate state reduces these probabilities still further. The external validity of an estimate of an investment return obtained in a single year is evidently extremely low in both agricultural environments.

In Appendix C we exploit the relatively large number of geographic clusters in the Ghana data to additionally assess intertemporal external validity across space. We explore to what extent estimates of parameters $\widehat{\boldsymbol{\alpha}}$ from one population can be used to characterize the intertemporal distributions of profits returns in other populations. We find that there clearly is information in the aggregate shocks that we measure that can be used to improve estimates of the returns to investment outside a given sample. However, there are cases in which the model predicts very different returns than are realized in a cluster. We also assess how well our estimate of $\boldsymbol{\alpha}$ predicts out-of-sample investment returns across years in the clusters by restricting estimation to just two years. We find that the predicted returns are reasonably close to those observed in the out-of-sample year.

C. Can Spatial Variation in Macro Shocks be used to Identify α?

The estimates in Table 2 present a significant challenge to researchers interested in the returns to investment, new technologies, or market innovations in agriculture. An alternative to estimating $\boldsymbol{\alpha}$ from multiple observations of a single population over time could be to exploit cross-sectional variation in $\boldsymbol{s_1}$ over space at a given time. Often, experimental designs include interventions across multiple administrative units spread across space. The data in northern Ghana encompass a sufficiently broad geographic area (150 km across its greatest extent) and rainfall realizations are sufficiently localized so that there is intra-annual variation in weather realizations across communities. Therefore, it may be possible to estimate the returns to planting stage investment and the interactions of those returns with weather realizations in a single cross-section. Using a cross section requires abandoning the fixed effects specification. This method would thus only be appropriate if the realization of $\boldsymbol{s_1}$ is uncorrelated with fixed unobservable characteristics that might affect net profits. The issue is how important these correlations are.

The results from estimating the relationship between rainfall realizations and returns to investment (evaluated at the median level of investment and median total rainfall) using each of the three cross-sectional waves of the panel are presented in Table 4.[15] Figures 6a and 6b report, respectively, the relationship between investment returns by the rainfall index, conditional on the normalized mean value of total rainfall, and by total rainfall, conditional on the normalized mean value of the index, for all three survey years. As can be seen, all three plotted relationships for the index measure lie within the confidence intervals of the panel-based relationship. However, for the seasonal rainfall measure, only one of the three cross-sectional plots happens to lie inside the panel-based confidence intervals. These results thus suggest that while it may be possible that cross-sectional rainfall variation in a single year can proxy for intertemporal variation, it is not possible to know, and we see at least two cases (of three) in which the cross-sectional relationships for one macro shock measure deviate importantly from the panel estimates.

As a consequence, the strategy of using cross-sectional variation in the shock to identify α, and pairing this estimate with observations of shocks realized later to estimate the returns to investment in years following a survey is problematic. To assess this, we use estimates $\widehat{\boldsymbol{\alpha}}_{2009}$ from the 2009 cross-section to predict cluster-year returns in 2010 and 2011, and $\widehat{\boldsymbol{\alpha}}_{2010}$ to predict cluster-year returns in 2011. The results are erratic. The correlation of returns we observe over the 11 clusters of villages in 2010 with those predicted using the 2009 cross-section (and the 2010 rainfall realizations) is -0.20 ($p = 0.56$); the correlation between observed returns in 2011 with those predicted from the 2009 cross-section is 0.38 ($p = 0.25$). The correlation between the 2011 observed returns and those predicted using the 2010 cross-section (with 2011 rainfall realizations) is 0.44 ($p = 0.18$).

[15] These cross-section estimates include a rich set of land and household characteristics in place of the fixed effects used in the panel. The cross-sectional regressions condition on 16 categories of age/sex/education of household members, three measures of household head education, household wealth at the baseline survey, seven soil types, five topographic categories, three erosion status indicators, four land tenure indicators, tree cover and plot distance.

Cross-sectional variation in the rainfall distribution generates biased estimates of $\boldsymbol{\alpha}$ when it is correlated with unobserved determinants of investment returns. We illustrate the salience of this concern by examining the relationship between per-capita wealth and rainfall in Appendix D. Because wealth may affect returns through at least two channels – farmers with higher wealth may be more willing to undertake riskier investments (Rosenzweig and Binswanger, 1993) or face less binding liquidity constraints – profits will be more sensitive to rainfall in the cross-section if wealth is associated with more risk-taking; if the latter dominates, returns to investments will be lower where rainfall is ample if there are diminishing returns to investments. We find that lagged wealth predicts rainfall realizations differentially across years, suggesting both why cross-sectional estimates of $\hat{\boldsymbol{\alpha}}$ are inconsistent and why they perform erratically in predicting investments returns across years.

II. SRI LANKA MICRO ENTERPRISES 2005-2007

One might think that rainfall variability, intra-seasonal and inter-annual, is only relevant to farming. However, the agricultural sector in most low-income countries employs a large fraction of the labor force. Fluctuations in farm income thus may affect the demand for non-agricultural products. In this section we show that the returns to investment in non-farm micro-enterprises in Sri Lanka, where 35% of the labor force is employed in agriculture, vary significantly within the year and correspond to the local seasonal agricultural cycle.

De Mel *at al.* (2009) carried out a panel survey of 408 non-farm micro-enterprises (less than $1,000 in invested capital) in three southern and southwestern districts of Sri Lanka for nine consecutive quarters between 2005 and 2007. The aim of the study was to estimate the returns to investment by implementing an RCT in which cash and capital equipment were offered to a randomized subset of firms after the first and third survey rounds. We use their data to explore whether the issue of returns variability due to external aggregate shocks is salient outside of the agricultural sector. While in general it may be difficult to isolate and measure the source of demand shocks for non-agricultural enterprises, in most low-income countries, including Sri Lanka, a large fraction of income is derived from the agricultural sector. And, a salient feature of agricultural production is that it is seasonal. Thus, to the extent that the Sri Lanka microenterprises serve local populations, it is not unlikely that demand is linked to the seasonal patterns of agricultural incomes.

The authors of the study pooled the quarterly profit data and provided one estimate of capital returns for the whole survey period. They believed that the observed inter-quarter profit variability was predominantly measurement error, and indeed they trimmed the top 0.5% of quarter-to-quarter differences.[16] We use their trimmed data and their IV strategy, which is similar to that we employed in our analysis of investment returns among farmers in Ghana. That is, we used the randomized interventions as instruments to estimate investment returns, but we allow the returns to differ by quarter to allow for a seasonal pattern of returns and to test whether the seasonal patterns, if any, are consistent with the agricultural calendar.

A component of the survey instrument sheds light on the issue of whether the measured profit variability in the data is due to external aggregate demand shocks or was simply measurement error. The entrepreneurs were asked to provide information on their expectations about their enterprise's profits in the subsequent quarter. Specifically, they were asked to indicate the lowest amount of monthly profits they thought they could earn three months from the survey date and the highest. The answers indicated that the respondents anticipated significant real quarterly profit variability. We computed the triangular coefficient of variation based on the answers, which was 0.5. Interestingly, the average coefficient of variation across firms for measured quarterly profits over the nine quarters was also 0.5.

More interestingly, the entrepreneurs were also asked to name the most important source of variability in profits. While 42.9% gave illness as the most important reason (own or family), 32% indicated that the most important factor causing a deviation in profits was an aggregate shock to product demand, including weather variation (8.2%), changes in income among customers (14.0%), and disruptions of infrastructure (road closing, power outages), 10.3%.

The first column of Table 5 reports our enterprise and quarter fixed effects IV estimates of the return on investment from the enterprise data that pools the profit data across quarters,

---

[16] Note that if their assumption is true, then the trimmed data merely have smaller sampling variability, but if trimming excludes important firms then our results will understate total intertemporal variation.

as in the original study. These estimates are obtained from the same unbalanced panel of 385 firms with at least three rounds of information as used by de Mel *at al.* Our specification differs from theirs in two ways, however. First, we allow the investment responses to the treatments to vary by year and by the gender of the entrepreneur, as in subsequent work the same authors found that treatment response was significantly greater for male respondents (de Mel *et al.*, 2012). Second, given the reports by the respondents on the importance of illness in influencing profits, we include a dummy variable indicating whether or not the entrepreneur reported any illness that resulted in his or her losing days of work. The estimate of illness variable coefficient is statistically significant, and indicates that when the entrepreneur is ill profits are lower by 10%, showing an important supply shock to profits. The return on investment is also statistically significant, using the conventional standard error as our criterion, and is similar in magnitude to that found in the original study.

In the second column, we report estimates from a specification that allows the returns to vary by quarter. In all quarters returns are positive and statistically different from zero. The point estimates, however, vary from .066 to .133, a factor of 100%. The difference between the lowest quarterly return and the highest is statistically significant. However, we cannot reject the hypothesis by conventional standards that the returns are equal across *all* quarters.

Because every quarter of the year is represented twice in the data we can test for the existence of a seasonal pattern in returns based on the quarterly returns estimates and assess if that pattern is related to agricultural income seasonality. This would not be possible if profits were measured over a short time period (say net revenue in the last week) and interview dates varied widely across quarters for the same firm. However, the window for revenue is the past month and the data set provides the exact date of the firm interview for all but one of the rounds. The interview date information indicates that 99.0% of the firms were interviewed in the same month in each quarter, and the inter-year differences in quarter-specific interview dates were no more than two weeks for 92.7% of respondents. Thus, the reference period is comparable within a quarter across years for firms and quarterly patterns are likely identifiable if they exist.

In the three districts in which the sampled microenterprises were located, the agricultural dry season is from December to March. If demand is driving returns we should expect to see that in the dry quarter the difference in investment returns between the treatment and control group should be lower than in other quarters. The quarterly returns estimates in the second column of Table 5 indicate that we can indeed reject the hypothesis of no seasonal pattern to investment returns (similarity of returns across seasons). We thus obtained new estimates of investment returns that are permitted to vary by season rather than by quarter. These estimates are reported in the first column of Table 6. For this specification, all the seasonal investment returns are significantly different from zero and we can also reject the hypothesis that the returns do not differ across seasons. There exists real seasonal returns variability. Moreover, this variability appears to be driven by the seasonal pattern of rainfall - the lowest return is in the January-March quarter, corresponding to the area-specific dry season. The point estimates indicate that had the experiment taken place only in that quarter rather than any of the other quarters the return estimate would have been on average 25% lower. Moreover, the pooled estimate of the returns to investments, net of season and entrepreneur fixed effects (column 2), .098, is not equal to the mean of the seasonal estimates of .087.

III. RETURNS TO SCHOOLING: URBAN INDONESIA 1993-2000

In this section we examine whether short-term aggregate shocks affect the estimated rate of return to education using panel data on urban wage workers in Indonesia. Schooling decisions, like enterprise and farm decisions, are made before the realization of the stochastic shocks relevant to future earnings streams and schooling decisions. Therefore, these decisions will be made based on expected returns, and rates of return to schooling are also subject to macro shocks. Unlike for firm and farm investments, however, investments in schooling have payoffs over a long period and schooling investments are not regularly repeated year after year. The intertemporal external validity issue for schooling returns is thus not so much that the year-to-year variability in schooling returns affects schooling decisions but how informative a single-year's contemporaneous estimate of the return may be to those making schooling decisions.[17]

[17] Jensen (2010), for example, presents a single year estimate of the skilled-unskilled wage differential to households making schooling decisions to test whether lack of information on schooling returns is a factor explaining low schooling. How the respondents responded to the estimate would depend on their knowledge of how informative a single year's estimate of the schooling return was for gaging the expected return, which would in turn depend on the realization of the shock in that year relative to its expected value.

As noted above, there have been studies that have examined fluctuations in returns in developed countries. In those studies the sources of the fluctuations are not identified and are typically measured by endogenous outcomes such as unemployment, and thus they are not useful for assessing the external validity of any given return estimate. Naturally, how the shock affects the investment return depends on the source of the shock.

Many developing countries are less diversified than developed countries in terms of exports, and incomes are thus more sensitive to fluctuations in individual commodity world prices. We study Indonesia because it is a country whose incomes are heavily dependent on the value of its oil exports and thus on shocks to world oil prices. In addition, Indonesia also has experienced external shocks to the value of its currency. We consider here not only whether but how both these shocks affect single-year estimates of schooling returns by exploiting the cross-sectional variation in the composition of industry.

The realizations of macroeconomic shocks are the same across all provinces, but they may have varying impacts on the distribution of wages and thus possibly schooling returns across provinces because of cross-province differences in industry composition. For oil price shocks, the effect in the short run on schooling returns in a given area, absent perfect mobility, will depend on the share of the workforce in the petroleum sector and the skill-intensity of the petroleum industry. For exchange rate shocks we expect that the wage premium for schooling increases more in provinces with a higher share of high skilled employment in open-to-trade industries when the exchange rate depreciates. Thus, while in the long run the expected value of the wage premium is equalized across provinces, the slope of the relationship between the realized wage premium in a single year and these macroeconomic shocks varies according to the industrial structure of the province. These short-run effects identified off province differences mimic the longer-run aggregate effects of oil price and exchange rate trends at the country level to the extent that Indonesia is a closed economy with respect to labor and capital.

To assess to what extent and how macro shocks affect estimates of the returns to schooling we use panel data from the first three rounds of the Indonesia Family Life Survey (IFLS), focusing on short-term contemporaneous variation in macro shocks, as we did for agricultural and non-agricultural investments. We selected the IFLS, a panel survey based on a representative sample of households in 13 major provinces in Indonesia begun in 1993, for three reasons: First, we can construct an annual hourly wage series for individual (wage and salary) workers from 1993-2000. The survey rounds were in 1993, 1997, and 2000, but every wave has a retrospective earnings history so we can fill in the inter-survey years using the wage and employment histories.

A second reason for selecting the IFLS for this time period is that Indonesia was affected by two external macroeconomic shocks. First, during this period there were large fluctuations in the world price of oil. Second, the financial crisis of 1998 substantially altered the value of Indonesian currency on the world market, potentially affecting the workforce in industries engaged in international trade relative to those in industries producing non-tradeables. Figure 7 shows the movements in the price of oil and the rupiah exchange rate from 1990 to 2000, both of which display considerable volatility over the period. We can assess to what extent these macro shocks affected the returns to schooling in urban areas in the eight years of the panel.[18]

The third reason for selecting the IFLS data is that, with our estimates of the sensitivity of schooling returns to oil price and rupiah variation in the decade of the 1990's, we can assess the external validity of the Duflo (2001) estimates of both the effects of school building on schooling attainment and the returns to schooling, based on a school-building initiative in the 1973-78 period and earnings measured from the 1995 Indonesia census. We do this in the next section.

In our panel sample of male urban wage and salary workers aged 16-80 with eight years of non-zero earnings from the IFLS, the average coefficient of variation in real earnings is a nontrivial 0.43. This value is comparable to our estimate of micro-enterprise profit variation from the Sri Lanka panel.[19] The main issue is again whether the returns to schooling vary significantly across years, and to what extent this variation is attributable to aggregate shocks.

---

[18]There have been studies of the effects of oil price shocks on the US economy. These have shown that from 1948 to 1972, increases in the world price of oil played a significant role in US recessions (Hamilton, 1983) but had no detectable adverse effects after 1973 (Hooker, 1996). These results are not obviously applicable to Indonesia, as the petrochemical industry constitutes a much smaller share of industry employment in the United States than in Indonesia. Additionally, governmental revenues in the US are not nearly as dependent on oil revenues.

[19] Some of this variation is also due to measurement error. In these data we can obtain an upper bound on measurement error by comparing the overlap of contemporaneous wage reports and retrospective wage reports from a subsequent survey round (for example, in the 1997 round, respondents report their wage in 1993, the preceding survey year). We find that the measurement error in wage rates accounts for approximately 40% of the cross-sectional variance, assuming independence of errors. But it is unlikely that there is as much measurement error in earnings as there is in agricultural or micro-enterprise profits computed from survey data, where we found evidence of systematic factors that altered investment returns.

The first column of Table 7 reports the OLS estimate of “the” rate of return to schooling based on pooling all observations for urban male wage and salary workers with at least two years of non-zero earnings over the eight-year panel. The specification is the standard “Mincer” log wage specification, including schooling years, age and age squared. We also include year fixed effects. The set of individual linear year coefficients are statistically significant, and indicate significant variation in hourly wage levels across years. In the next column, the rates of return are allowed to vary by year. All of the year-specific schooling returns are statistically different from zero by conventional standards, and we can also reject the hypothesis that all the year-specific returns are equal.

The OLS returns estimates in the first two columns may be biased due to the endogeneity of schooling. In the third column of Table 7 we report worker fixed effects estimates, which eliminates the influence of any worker characteristics that jointly affect schooling and earnings. In this specification we cannot identify the returns to schooling, but we can obtain consistent estimates of differences in returns across years. The year-specific returns coefficients in the third column are relative to that in the omitted year, 2000. These estimates too lead to rejection of the hypothesis that the schooling returns are equal across years. The point estimates indicate, for example, that the schooling returns in 1998 and 1999 are from 2.6 to 2.7 percentage points, respectively, below that for 2000. These differences are statistically significant

Just as for the agricultural samples, we are interested in understanding the sources of variation in returns. We not only want to estimate α’s for schooling returns but also to explore the mechanisms. We thus examine the influence on the return to schooling investment of variation in the world price of oil over the sample period and of changes in the rupiah exchange rate in the last three years of the panel, when we think the variation in the value of the rupiah is mainly due to forces external to Indonesia rather than to domestic factors.

With respect to the oil price, as noted, whether an increase in the price of oil increases or decreases the schooling return in the short run depends on whether it raises the demand for skilled relative to unskilled labor. We would expect that an increase in the oil price would likely raise the returns to schooling. First, an increase in the price of oil would lead to an expansion in the demand for workers in the petrochemical industry, and the 1990 Indonesian Census indicates that workers in the petrochemical industry are more skilled relative to the overall workforce, whether measured by mean years of schooling or proportion of workers with secondary education. Second, because government revenues depend heavily on the price of oil in Indonesia, we would expect that employment in the public sector would also expand, and generally public-sector employees are more skilled. For example, the INPRES school-building program in 1973-78, financed with the new revenues resulting from an increase in the price of oil in that period, increased as well the employment of teachers, who are significantly more educated than the average Indonesian worker. We thus would expect that in the short run, before workers can re-allocate among industries, a rise in the price of oil will benefit skilled relative to less skilled workers and raise the return to schooling in areas with a larger employment of oil workers.

We re-estimate the log wage equation omitting the set of individual year-schooling interactions and replacing them with an interaction between the oil price and schooling to estimate α for oil price variation. Because the oil price is time-varying, we can identify the influence of shocks to oil prices on schooling returns while including the worker fixed effect.[20] Column 1 of Table 8 reports the OLS estimates of the return to schooling and of α. The specification includes schooling, age, age squared, the interactions of schooling and the age variables with the oil price, the oil price, and year. As expected α is positive and statistically significant, while the direct effect of the price of oil on wages (for the unskilled) is actually negative. When the worker fixed effect is included in column 2, the estimate of the oil price α retains its statistical significance, and increases in magnitude by 50%, while the direct oil price effect is slightly attenuated.

The point estimate of α in the second column implies that a one standard deviation (measured over the decade) increase in the price of oil would modestly raise the return to schooling, by a statistically significant 0.7 percentage points. However, the estimate also implies that the rise in the price of oil from 1998 to 2000 ($16 per barrel) increased the return to schooling by 2.5 percentage points, accounting for all of the observed increase in the return to schooling in that interval seen in column two of Table 7.

[20] 2.7% of respondents moved to another province over the eight-year period. These respondents were dropped. In the fixed effects specification, only if such movers had schooling returns unusually sensitive to the oil price shocks will this affect (attenuate) our estimates of the shocks on schooling returns.

We have assumed that the oil price increased the relative demand for skilled versus unskilled workers in Indonesia as the reason why we found the oil price α to be positive. Next we more directly test the hypothesis that changes in schooling returns due to macro shocks arise from the effects they have on the relative demand for skilled and unskilled labor, exploiting the movements in the exchange rate value of the rupiah over the 1997-200 period. As seen in Figure 7, the exchange rate increased threefold between 1997 and 1998 due to the worldwide financial crisis in 1998. To carry out the analysis, we make use of a method employed in the trade literature to examine the effects of changes in trade policies on schooling choice. The key idea is that the policies (e.g., tariff reforms) affect the returns to schooling via their effects on industries differentiated by skill intensity and sensitivity to the policy as measured by trade openness (e.g., Edmonds *et al.*, 2010; Atkin, 2015). In this literature, the estimates are of the effects of the shifts in trade variables via industry structure on schooling choice. Here we estimate the effects of the external trade shock on schooling returns, which influence the decisions of agents making schooling choices. Analogous to the inquiries in the trade literature, we wish to test whether in areas (provinces) where there are more workers employed in industries that are both more open, and thus disproportionally affected by a change in the value of the rupiah, and more skilled, an exogenous rise in the exchange rate raises the returns to schooling. That is, the more open are skilled industries in an area, the greater the increase in the return to schooling.

To carry out this analysis we constructed a standard measure of trade openness by industry ((exports + imports)/ value added) O_k, where *k* indexes industry.[21] We also obtained information on the skill-intensity (share of workers with secondary schooling) of industries S_k from the 1990 Indonesia Census. From these data we created our average skill/openness measure by province: $SO_j = \sum_k O_k S_k G_{jk}$. We then interacted that measure by the year-specific rupiah exchange rate xr_t and the schooling of the respondent. Our hypothesis is that the coefficient on the triple interaction between (provincial) openness, (annual) exchange rates and (individual) education ($SO_j \cdot xr_t \cdot a_{0i}$) is positive, implying that a rise in the exchange rate

[21] We obtained industry-specific information on imports and exports in 1990 for Indonesian industries from the UN ComTrade database and industry output and province- and industry-specific GDP share for each province *j* G_{jk} from Statistics Indonesia
http://comtrade.un.org/db/default.aspx and https://www.bps.go.id/linkTabelStatis/view/id/1200.

increases the returns to schooling more in areas in which there is a heavier concentration of industries that are open and skill-intensive.

The last two columns of Table 8 report the OLS and fixed effects estimates of the log wage equation, based on the years 1997-2000, including the oil price variables, the exchange rate and openness/skill variables. As can be seen the returns to schooling do rise when the exchange rate increases in areas with more open and skilled industries. Increases in the oil price also increase the schooling rate of return in this sub-period. The point estimate from the individual fixed-effect specification indicates that at the mean of the *SO* measure of openness/skill for Indonesia, the large rise in the exchange rate from its value in 1997 to its average value in 1998-2000 after the financial crisis increased the rate of return to schooling by 2.4 percentage points. If we use Duflo's (2001) estimate of the mean return level for the decade, which we show to be externally valid below, that represents an increase of over a third in the schooling return. Clearly, domestic schooling returns are sensitive to global macro shocks in Indonesia. While exchange rate volatility in Indonesia after 2000 seems to be relatively low, as can be seen in Figure 8 oil price volatility in the first 15 years of the 21st century was substantially higher than it was in the 1990's. We would therefore expect that investments in schooling have become much riskier, and single-year estimates of rates of return to schooling are now less reliable indicators of the expected returns to schooling.

A. Assessing *Ex Post* the External Validity of Duflo's (2000) Estimates of School Access and Schooling Returns.

Our findings imply that estimates of the sensitivity of returns to aggregate shocks may be necessary to more fully assess the external validity of estimates, and this most likely entails obtaining estimates of returns over multiple years. In this section, we show that attention to aggregate shocks can shed light on the external validity of estimates even for projects that obtain estimates from only a single year. In particular, our model and estimate of the effect of the oil price on returns to education (α) allow us to *ex post* evaluate the external validity of Duflo's (2001) estimates of both the effects of increasing access to schools via school building on schooling attainment and the returns to schooling. In that study, the massive program building schools in Indonesia in the years 1973 through 1978 was used to estimate the effects of increasing school access on schooling attainment and the consequences of the increased schooling for earnings, based on the earnings for the relevant cohorts measured in the single period 1994-1995.

Figure 8, which displays the changes in oil prices over the period 1960 through 2016, highlights the prices for years relevant to the Duflo study. What is apparent is that the first year of the school building initiative was marked by a permanent increase in oil prices. This fact is not surprising, as the financing of the program was supported by the increase in oil revenues accruing to the government. If the jump in oil prices at the initiation of the program increased expectations of future levels of oil prices, given our estimate of a positive α, the oil price increase alone would have increased schooling attainment by altering the expected rate of return to schooling even if access to schools had remained unchanged. If both the control group – those cohorts who chose schooling before the program was put in place - and the treatment group equally changed their expectations of schooling returns, the estimated impact of the program has internal validity. However, as shown in expression (14), the estimated effect of the program, given our finding that $\alpha>0$, while indeed causal, likely overstates the effect of increasing school access that would occur when returns expectations are stable.

What about the external validity of the estimate of the rate of return to schooling obtained by Duflo? The return estimate was obtained based on earnings measured in a single year, from the 1995 Census. An issue is what is the appropriate return to compare Duflo's single-year estimate? One benchmark might be the expected or average return over the decade, say the return associated with the decade average oil price. The average oil price from 1991-2000 was \$18.2, while the price of oil in the Census earnings reference year used by Duflo was between \$15.5 and \$16.9, a difference of only about \$2 from the mean price. Given our estimate of α, this difference implies that Duflo's estimate of the expected return over the decade was too low by only 0.3 percentage points - 1995 happened to be a typical year. In contrast, if Duflo had used 2000 as the reference year (say, using the 2000 Indonesian Census), when the oil price had spiked to \$27.6, our estimate of α implies that her estimate of the schooling return would have been too high by about 1.5 percentage points, 21% above the expected value for the decade, given her externally valid estimate of the schooling return of 7%.

IV. CONCLUSION

While it is known that aggregate shocks can affect micro parameter estimates and their confidence intervals, there is little evidence of the empirical importance of the existence of time-varying, aggregate shocks for drawing inferences about the external validity of internally-valid causal estimates that are typically obtained in one time-period. The key set of parameters needed to assess temporal external validity is that describing how aggregate shocks affect the returns to the actions, α. Estimates of shock effects on returns, however, are rarely obtained in existing studies, even in those studies using information from multiple years, which could in principle be used to identify α. Because it is necessary to have information on not only the sensitivity of investment returns to macro shocks but also the parameters of the relevant shock distributions most existing studies therefore also do not provide appropriate parameter confidence intervals to the extent that aggregate shocks are important. These issues would appear to be particularly salient for studies from developing countries, where a large fraction of the population is employed in agriculture and therefore subject to the vagaries of both weather and global commodity price shocks.

In our empirical investigation of the sensitivity of estimates of investment returns to external aggregate shocks using data from four low-income countries we have found evidence that investment returns, whether returns to agricultural investments, non-agricultural enterprise investments or urban investments in human capital, differ significantly from year to year. And, where we could measure the impact of specific aggregate shocks, such as for agriculture and for schooling, we found that the α's were large in magnitude and in sign consistent with economic theory. Moreover, where we could also reliably estimate the parameters of the distribution of one key aggregate shock, rainfall in India and Ghana, we showed that the probability of a single year estimate being within a reasonable interval around the expected investment return are very low. Thus, for rain-fed agriculture, single-year estimates of the returns to investments or interventions have extremely low temporal external validity. Our estimates also imply that our knowledge of the impact of investments or policy interventions may be far less certain than existing estimates of confidence intervals imply. Agents making investment decisions and policy-makers making policy decisions in a stochastic world face substantially more risk than measured by conventional estimates that ignore aggregate shocks.

We also showed that in one data set where we had cross-sectional variation in rainfall, estimates based solely on cross-sectional variation did not reliably replicate those obtained using time-series variation that enable control for spatial and agent fixed effects. This is not unexpected, as draws from a stationary distribution will be correlated with the parameters of the distribution. In the case of agriculture and rainfall, we found that farmer wealth levels and rainfall means were positively correlated, such that the cross-sectional estimates of the sensitivity of output to rainfall obtained in Karlan *et al*. (2014) were positively biased, as would be expected if wealth matters for the level and/or the risk properties of investment decisions.

What is the remedy for enhancing temporal external validity and assessing the external validity of estimates in a world with stochastic shocks? First, micro models of individual decision-making need to identify and incorporate the role of aggregate shocks. Theory should suggest which states of nature matter, and how, for the outcomes and interventions of interest as well as what information agents have about future shocks when they make decisions. Second, information on the time-series of the relevant states of nature (shocks) for the population or site being studied should be obtained along with information on the realization of the shocks for the study period. Ideally, identification of α should also become a component of any well-designed study. This process likely requires information spanning multiple periods, as cross-sectional shock variation can be correlated with site-specific unobservables that influence outcomes, as we have shown.

Typically, though, an analyst using data from one particular context may not have the data available to estimate α. Knowledge of the context, of agents' decision-making processes and relevant technologies, however, can generate a prior belief regarding its sign and magnitude. These, coupled with estimates from external sources of the distribution of the relevant aggregate shocks, as we have shown for Duflo's (2000) estimates of schooling returns, can guide informed speculation *ex post* regarding the relationship between any single-period estimate of returns to expected returns or the returns from future actions.

REFERENCES

Allcott, Hunt. (2015). “Site Selection Bias in Program Evaluation.” *Quarterly Journal of Economics* 130(3): 1117-1165.

Attanasio, O. and B. Augsburg (forthcoming). “Holy Cows or Cash Cows.” *Economic Development and Cultural Change.*

Donald W. K. Andrews. "Cross-Section Regression with Common Shocks." *Econometrica* 73, no. 5 (2005): 1551-585.

Angrist, Joshua and Alan B. Krueger. (1991). “Does Compulsory School Attendance Affect Schooling and Earnings?” *The Quarterly Journal of Economics*, 106(4), 979-1014.

Angrist, Joshua and Alan B. Krueger. (1992). “Estimating the Payoff to Schooling Using the Vietnam-Era Draft Lottery.” NBER Working Paper No. 4067.

Atkin, David. (2015). “Endogenous Skill Acquisition and Export Manufacturing in Mexico.” NBER Working Paper No. 18266.

Bandiera, Oriana, Robin Burgess, Narayan Das, Selim Gulesci, Imran Rasul and Munshi Sulaiman. (forthcoming). “Labor Markets and Poverty in Village Economies.” *Quarterly Journal of Economics*.

Banerjee, Abhijit, Sylvain Chassang, and Erik Snowberg. (2016). “Decision Theoretic Approaches to Experiment Design and External Validity.” NBER Working Paper No. 22167.

Banerjee, Abhijit, and Esther Duflo (2008). “Do Firms Want to Borrow More: Testing Credit Constraints Using a Targeted Lending Program,” BREAD Working Paper No. 005, 2008.

Banerjee, Abhijit, Esther Duflo, Rachel Glennerster and Cynthia Kinnan. (2015). “The miracle of microfinance? Evidence from a Randomized Evaluation,” *American Economic Journal: Applied Economics* 7(1): 22-53.

Banerjee, Abhijit, Esther Duflo, Nathanael Goldberg, Dean Karlan, Robert Osei, William Parienté, Jeremy Shapiro, Bram Thuysbaert, and Christopher Udry. 2015. “A Multi-faceted Program Causes Lasting Progress for the Very Poor: Evidence from Six Countries.” *Science* 348 (6236): 1260799-1–1260799-16. doi:10.1126/science.1260799.

Beegle, K., J. Goldberg & E. Galasso (forthcoming). "Direct and Indirect Effects of Malawi`s Public Works Program on Food Security." *Journal of Development Economics*.

Bils, Mark and Klenow, Peter J. (2000). “Does Schooling Cause Growth?” *American Economic Review* 90(5): 1160–83.

Bloom, Nicholas, Benn Eifert, Aprajit Mahajan, David McKenzie, and John Roberts. (2013). "Does Management Matter? Evidence from India," *Quarterly Journal of Economics* 128: 1-51.

Cai, Hongbin, Yuyu Chen, Hanming Fang, and Li-An Zhou. (2015). “The Effect of Microinsurance on Economic Activities: Evidence from a Randomized Natural Field Experiment.” *Review of Economic and Statistics* 97(2): 287-300.

Card, David. (1995). “Using Geographic Variation in College Proximity to Estimate the Return to Schooling.” *Aspects of Labour Economics: Essays in Honour of John Vanderkamp*, Louis Christofides, E. Kenneth Grant and Robert Swindinsky, eds. University of Toronto Press.

Caselli, Francesco and Wilbur John Coleman II. (2006). “The World Technology Frontier.” *American Economic Review* 96(5): 499-520.

Dehejia, Rajeev, Cristian Pop-Eleches, and Cyrus Samii. (2016). “From Local to Global: External Validity in a Fertility Natural Experiment.” NBER Working Paper No. 21459.

de Mel, Suresh, David McKenzie, and Christopher Woodruff. (2008). "Returns to Capital in Microenterprises: Evidence from a Field Experiment," *The Quarterly Journal of Economics*, 123(4), 1329-1372, November.

de Mel, Suresh, David McKenzie, and Christopher Woodruff. (2012). “One Time Transfers of Cash or Capital Have Long Lasting Effects on Microenterprises in Sri Lanka.” *Science* 335.

Deaton, Angus and Nancy Cartwright. (2016). “Understanding and Misunderstanding Randomized Control Trials.” NBER Working Paper 22595.

Di Pietro, Georgio and Lucas Pedace. (2008). “Changes in the Returns to Education in Argentina.” *Journal of Applied Economics* 11(2): 259-279.

Duflo, Esther. (2001). “Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment.” American Economic Review 91(4).

Duflo, Esther, Michael Kremer, and Jonathan Robinson. (2011). "Nudging Farmers to Use Fertilizer: Theory and Experimental Evidence from Kenya," *American Economic Review*, 101(6).

Edmonds, Eric V., Nina Pavcnik and Petia Topalova. (2010). “Trade Adjustment and Human Capital Investments: Evidence from Indian Tariff Reform.” *American Economic Journal: Applied Economics* 2(4): 42-75.

Efron, Bradley. *Large-scale inference: empirical Bayes methods for estimation, testing, and prediction*. Vol. 1. Cambridge University Press, 2012.

Foster, Andrew D. and Mark R. Rosenzweig. (1995). “Learning by Doing and Learning from Others: Human Capital and Technical Change in Agriculture.” *Journal of Political Economy* 103(6).

Giné, Xavier, and Dean Yang. (2009). “Insurance, Credit, and Technology Adoption: Field Experimental Evidence from Malawi.” *Journal of Development Economics* 89 (1) (May): 1–11.

Hamilton, James D. (1983). “Oil and the Macroeconomy since World War II.” *Journal of Political Economy* 91(2): 228-248.

Hahn, Jinyong, Guido Kuersteiner and Maurizio Mazzocco. 2016. “Estimation with Aggregate Shocks.” Manuscript, UCLA.

Hanna, Rema, Sendhil Mullainathan Joshua Schwartzstein. (2014). “Learning Through Noticing: Theory and Evidence from a Field Experiment.” *The Quarterly Journal of Economics.*

Hooker, Mark A. (1996). “What happened to the oil price-macroeconomy relationship?” *Journal of Monetary Economics* 38:195-213.

Huffman, G.J., E.F. Stocker, D.T. Bolvin, E.J. Nelkin, R.F. Adler, 2012, last updated 2013: TRMM Version 7 3B42 and 3B43 Data Sets. NASA/GSFC, Greenbelt, MD. Data set accessed at

James, William, and Charles Stein. “Estimation with Quadratic Loss.” In *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability*, 1:361–379, 1961.

Jensen, Robert. (2010). “The (Perceived) Returns to Education and the Demand for Schooling.” *The Quarterly Journal of Economics*. 515-548.

Karlan, Dean, Isaac Osei-Akoto, Robert Osei, and Chris Udry. (2014). “Agricultural Decisions after Relaxing Credit and Risk Constraints,” *Quarterly Journal of Economics*, 129(2).

Krueger, Alan B. (1993). “How Computers Have Changed the Wage Structure: Evidence from Microdata, 1984-1989*.” The Quarterly Journal of Economics* 108(1): 33-60.

Lemieux, Thomas and David Card. (2001). “Education, Earnings, and the 'Canadian G.I. Bill’.” *The Canadian Journal of Economics* 34(2): 313-344.

Manzanas, R., L.K. Amekudzi, K. Preko, S. Herrera, J.M. Gutierrez. (2014). "Precipitation variability and trends in Ghana: An intercomparison of observational and reanalysis products." Manuscript, Grupo de Meteorologıa. Instituto de Fısica de Cantabria, CSIC-Univ. de Cantabri.

Meager, Rachael. (2016). "Aggregating Distributional Treatment Effects: A Bayesian Hierarchical Analysis of the Microcredit Literature." Manuscript: MIT.

Miguel, Edward and Michael Kremer. (2004). "Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities." *Econometrica* 72(1): 159–217.

Morris, Carl N., and Martin Lysy. "Shrinkage Estimation in Multilevel Normal Models." *Statistical Science*, 2012, 115–134.

Oreopoulos, Philip, Till von Wachter, and Andrew Heisz. (2012). "The Short- and Long-Term Career Effects of Graduating in a Recession." *American Economic Journal: Applied Economics* 4(1): 1–29.

Psacharopoulos, George. (1994)."Returns to Investment in Education: A Global Update." *World Development* 22(9): 1325–43.

Rosenzweig, Mark and Hans P. Binswanger. (1993). "Wealth, Weather Risk and the Composition and Profitability of Agricultural Investments." *The Economic Journal* 103(416): 56-78.

Suri, Tavneet. (2011). "Selection and Comparative Advantage in Technology Adoption." *Econometrica* 79(1): 159-209.

Tjernström, E., M.R. Carter and P. Toledo. (2013). "Identifying the Impact Dynamics of a Small Farmer Development Scheme in Nicaragua." *American Journal of Agricultural Economics*.

Zhang, Warren. 2016. "Worm Wars: Do Mass Deworming Campaigns Really Improve Education?" Manuscript: Yale University.

Appendix A: Price Variation and the Returns to Investment

A specific aggregate state variable that will typically be important and typically measurable over a long time-period is the price of output. Our model shows that it is relatively straightforward to assess the influence of price variation on investment returns. Let $\boldsymbol{s_1} = (p, \tilde{\boldsymbol{s}})'$. Let $y_{i1} = pf(a_{i0}, \tilde{\boldsymbol{s}})$. Therefore

$$\beta_i(a_{i0}, \epsilon_i, \boldsymbol{s}_1) = p\epsilon_i \frac{\partial f(a_{i0}, \tilde{\boldsymbol{s}})}{\partial a_{i0}} - 1. \tag{A1}$$

Given data from a single realization of $\boldsymbol{s_1} = \boldsymbol{s}_{1t}$ we can estimate only

$$\hat{\beta}(\hat{a}_0, \boldsymbol{s}_{1t}) = p_t \int \epsilon \frac{\partial f(a_{i0}, \tilde{\boldsymbol{s}}_{\boldsymbol{t}})}{\partial a_{i0}} g(\epsilon) d\epsilon - 1 \tag{A2}$$

and we can observe $\alpha_p = \int \epsilon \frac{\partial f(a_{i0}, \tilde{\boldsymbol{s}}_{\boldsymbol{t}})}{\partial a_{i0}} g(\epsilon) d\epsilon = \frac{\hat{\beta}(\hat{a}_0, \boldsymbol{s}_{1t}) + 1}{p_t}$. The linearity of the effect of variations in the price of output on the marginal return to investment eases the interpretation of the effect of this dimension of variation, and the $\hat{\beta}\left(\hat{a}_0, \left(p_t, \tilde{\boldsymbol{s}}_{\boldsymbol{t}}\right)'\right)$ observed for any particular realization of

p_t is sufficient to imply the $\hat{\beta}\left(\hat{a}_0, (p, \tilde{\boldsymbol{s}}_{\boldsymbol{t}})'\right)$ for any other realization of the price (given the other $k-1$ dimensions of $\boldsymbol{s}$), with

$$\hat{\beta}(\hat{a}_0, (p, \tilde{\boldsymbol{s}}_{\boldsymbol{t}})') = \left(\hat{\beta}(\hat{a}_0, (p_t, \tilde{\boldsymbol{s}}_{\boldsymbol{t}})') + 1\right)\frac{p}{p_t} = \alpha_p \cdot p \quad \text{(A3)}$$

for all a.

If price variability is the main source of stochastic variation in second-period outcomes, we can thus use the time-series of prices to quantify *ex post* the variability in profits or earnings faced historically by agents.

We have time series of prices relevant to farmers in both northern Ghana and the ICRISAT India villages. In each country, our observed returns correspond to $\hat{\beta}\left(\hat{a}_0, \left(p_t, \tilde{\boldsymbol{s}}_{\boldsymbol{t}}\right)'\right)$ in (A3). Was the study period a time of particularly low or high crop prices? We use (A3) to calculate how the mean realized return to investment (at mean levels of investment, total rainfall and rainfall timing index, as appropriate for each country) would vary with a Laspeyres price index for the most important crops in each study area. Unlike for rainfall, however, we do not know the distribution of future output prices and developing a model of the dynamics of commodity prices is beyond the scope of this paper. Therefore, we do not compute the relevant confidence intervals for investment returns that incorporate stochastic price variability. Instead, we document the changes in returns to investment in Ghana and India that have occurred over the duration of the price data and display the time series in Figure A1.[22] As can be seen, there are signficant fluctuations in the returns soley due to price changes. The estimated return in the highest index price year is 50-70% above that in the lowest price year within the seven-year span of these data, for a fixed rainfall realization.

Appendix B: Estimation of Returns to Investment

Estimates of the return to investment in India rely on the assumption that news regarding the probability of good rainfall is excludable from the profit function. There are three primary concerns regarding this excludability assumption. The first is that if the reactions to forecasts by farmers depends on their soil characteristics, it must also be true that the impact of rainfall on

[22] Maize, rice and yam in northern Ghana, using real July prices, and cotton, pigeon pea, green gram and sorghum for India using October prices.

investment returns is affected by soil properties. Thus we must include these interactions with investments in the second-stage profit function and they should have power in explaining profit variability. The second concern is that conditional on our specific measures of realized rainfall, the forecast of total rainfall may be correlated with an unmeasured dimension of rainfall that matters for profits. We measure realized rainfall as the total amount of rainfall over the year and the total amount of rainfall over the monsoon, as the IMD long-range forecast is the prediction for the total amount of rainfall over the monsoon. Binswanger and Rosenzweig (1993) have shown that the monsoon onset date is a salient feature of rainfall for farm profits in India. However, in the ICRISAT data we find that conditional on even a subset of our measures of rainfall (monsoon rainfall), the IMD forecast of total monsoon rainfall is not correlated with the onset date. Note that the village-year effects capture all time-varying aggregate shocks.

A third concern is that because the rainfall forecast for a given year is common to everyone in a village through its effect on input demand, a forecast of good (bad) weather could raise (lower) input prices - particularly wages - in a village. In principle it is also possible that there could be policy interventions (changes in regulated grain prices, emergency agricultural interventions, *ex ante* efforts to provide relief). These village-specific changes correlated with the forecast could affect profits directly. As noted, the village-year fixed effects are included in the profit function to absorb these effects. A casualty of including village-year fixed effects is that the direct effects of rainfall and lagged rainfall on profits are not identified.[23]

[23]A further concern that would make the forecast non-excludable is that the increased planting-stage investments induced by a favorable forecast reduce the farmer's resources available for subsequent production stages. In the model this is ruled out by the implicit assumption of perfect credit markets *within* the relevant production cycle. The ICRISAT survey data enable us to carry out a global separability test similar to that of Benjamin (1992). The basic idea is that exogenous changes in the family labor force should not affect profits if all input markets are unconstrained. Illness has a large random component (net of the household fixed effect), and illness can affect the family's ability to supply labor. For the years 2005, 2006, 2010 and 2011 the ICRISAT survey elicited information on the number of days that adult family members were ill in the *kharif* season. Household fixed effect estimates obtained for the total sample of farmers and the farm households in the Maharashtra villages of the effect of the number of sick days on total labor days in the kharif season indicate that for each day an adult was sick almost a third of a day of on-farm family labor was lost. The estimate is $LabDays_{it} = \underset{(.09)}{-0.34} Sickdays_{it} + \underset{(.0004)}{.002}\ TotRain_{vt}$. If liquidity constraints limited the ability of the household to substitute hired labor to make up for family labor days lost, an increase in sick days should therefore decrease profits. However, FE-IV estimates of the profit function for the Maharashtra farmers including the number of adult sick days (not reported) indicate that we cannot reject the hypothesis of separability

The randomized assignment of farmers to index insurance and cash grant treatments in Ghana is the foundation of the identification of the returns to investment conditional on aggregate shocks. We rely on the assumption that the assignment to alternative treatments affects net income only through the choice of planting season investments. The primary concern that arises with respect to this assumption is that conditional on planting season investments, assignment to the different treatments could influence later-stage cultivation decisions and thus be correlated with net income. This objection will not apply for assignment to the cash grant treatment if liquidity constraints do not bind with respect to expenditures within the growing season. The results of Karlan *et al*. (2014) showing that these households were able to substantially increase average planting season investments upon assignment to free or reduced cost insurance without any infusion of additional capital suggest that within season liquidity constraints are not binding. The identification assumption with respect to the insurance treatments is that conditional on the level of planting season investment and rainfall realizations, post-planting cultivation decisions are independent of insurance. Identification is threatened, however, if there is sufficient flexibility in cultivation opportunities after the conclusion of planting season investment that farmer decisions might be influenced by insurance status (e.g., conditional on planting stage cultivation decisions and rainfall realizations, an insured farmer decides to replant after a late season drought spell while an uninsured farmer does not). This identification concern is mitigated by the fact that the crucial decisions guiding end-season cultivation activities are the planted area, choice of crop, planting density, and chemical application, all of which are completed by the end of the planting period. By the completion of planting season, farmers have accrued 90% of total non-family labor costs for the entire season.[24]

Appendix C: Assessing Spatial and Intertemporal External Validity

We exploit the relatively large number of geographic clusters in the Ghana data to additionally assess intertemporal external validity across space. We want to know if estimates of

- despite sick days evidently significantly reducing on-farm family labor supply, an increase in the number of adult sick days has no impact on profitability.

[24] Estimates from a narrower definition of "planting season" that excluded fertilizer expenses are available from the authors (and were provided in an earlier version of the paper). The results are qualitatively similar to those presented here.

parameters $\hat{\alpha}$ from one population can be used to characterize the intertemporal distributions of profits returns in other populations. To do this, we divide the Ghana sample into a sequence of training and test subsamples in order to quantify the extent to which it is possible to predict some of the observed variation in returns to capital across cluster years within northern Ghana, where the technology of agricultural production is plausibly similar, based on estimates from only a subset of clusters.

In each case i we use the data in the training sample clusters to estimate the parameters $\widehat{\boldsymbol{\alpha}}_i$ in the model as specified in column 2 of Table 3. We then use those estimates along with the observed weather realizations in the test sample cluster-years to predict the cluster-year returns to capital in the test sample: $\hat{\beta}^i_{vt} = \boldsymbol{s}'_{\boldsymbol{vt}}\widehat{\boldsymbol{\alpha}}$. This prediction of returns is then compared to $\hat{\beta}_{vt}$, the "actual" returns to capital in each of the cluster years in the test sample as measured by estimating (15). The cluster-year returns as predicted by the two rainfall realizations are correlated with the realized returns across clusters: the correlation over clusters is 0.40 when each test sample is a single cluster ($p = 0.22$); 0.31 when each test sample is two clusters ($p = 0.03$) and 0.20 ($p = 0.01$) when each test sample is three clusters. It is also possible to predict changes over time within clusters. In 6 of 11 cases with a single cluster in each test sample, the correlation between predicted and realized returns over the three years is positive ($p = 0.23$). When each test sample is two clusters, 31 of the 51 time series correlations are positive ($p = 0.03$). And with test samples of three clusters each, 96 of the 164 time series correlations are positive ($p = 0.01$). Figure A2 provides a summary account of the relationship between realized returns in each cluster year, and the average across the test sample predictions of the return in each cluster year. There clearly is information in the aggregate shocks that we measure that can be used to improve estimates of the returns to investment outside a given sample. However, there are cases in which the model predicts very different returns than are realized in a cluster – for example cluster 2, year 3. Evidently there are a host of macro shocks other than our two dimensions of rainfall that affect average returns.

The three years of data from northen Ghana permit us to also explore the external validity of our fixed-effect IV estimates of $\boldsymbol{\alpha}$ across time and the clusters. In particular, we use the first two years of data (2009-10) to estimate consistent fixed-effect IV estimates of $\boldsymbol{\alpha}$ (reported in column 3 of Table 3), and use these estimates along with 2011 rainfall realizations to predict rates of return to investment across the 11 geographic clusters. The correlation between these predictions and the observed rates of return is 0.39 (p=0.23). Information about the realization of macro shocks, when combined with consistent estimates of the responsiveness of returns to those shocks, can be useful in improving the temporal external validity of estimates of the returns to an investment.

Appendix D: Wealth and Rainfall Realizations

To first see if household wealth and the mean of the rainfall distribution are correlated, we regressed the log of per-capita household wealth from the 2009 baseline survey on the mean of the rainfall distribution (and its square) over the 30 clusters, based on the 16-year time series of seasonal cluster-specific rainfall from the TRMM data set. The estimates are reported in the first column of Appendix Table A1. The estimates clearly indicate that households in areas that on average have more rain also have higher levels of per-capita wealth. Is it possible to just use a single-year cross-section of rainfall outcomes to test for this potential source of bias? Of course, in any given year higher rain is more likely in an area with a higher mean rainfall. A potential placebo test therefore is to regress baseline wealth on subsequent rainfall. In columns two and three of Appendix Table A1 we report estimates of the effects of rainfall and its square in 2009, 2010 and 2011, respectively, on log wealth measured in the 2009 baseline survey. In all years, the estimates indicate that rainfall and wealth are not independent, and in two of the three years wealth and the single-year rainfall outcome are positively related; but in 2009 rainfall is negatively correlated with baseline wealth. This demonstrates again that a single cross-section, given high rainfall variability, may not be revealing about longer term relationships.[25]

[25] Karlan *et al*. (2014) make use of the cross-sectional variation in rainfall realizations to identify the effects of insurance on risk-taking. In particular, they compared the responsiveness of farm output to rainfall across farmers randomly assigned to receive free or subsidized insurance compared to a control group not offerered insurance. The idea was that farmers offerred insurance would take more risk, and this would be expressed as a higher sensivity of their output to rainfall. However, because this regression makes use of cross-sectional variation in rainfall outcomes over the first two years of the panel (2009 and 2010), which we have shown are correlated with farmer wealth, it is likely that the estimates are biased positively, as long as wealth increases investment responses to insurance via liquidity effects. Consistent with this, re-estimation of their specification with village fixed effects results in a reduction in the rainfall-insurance treatment interaction coefficient by over a third and a loss of statistical signficance.

Table 1
Farmer Fixed Effects IV Estimates of Planting-Stage Investments on Net Returns,
Four ICRISAT Villages, 2005-2011

Variable	(1)	(2)
Planting-stage investment	7.15	1.54
	(2.42)	(3.99)
Planting-stage investment squared x10^{-5}	-7.26	-3.91
	(3.13)	(4.50)
Planting-stage investment x *Kharif* rainfall	-	0.0163
		(0.00639)
Planting-stage investment squared x *Kharif* rainfall x10^{-5}	-	-0.0123
		(0.00658)
Planting-stage investment x *Kharif* rainfall x depth 1	-	-0.0103
		(0.00358)
Planting-stage investment x *Kharif* rainfall x depth 2	-	-0.00259
		(0.00456)
Planting-stage investment x *Kharif* rainfall x sandy soil	-	0.00246
		(0.00330)
Planting-stage investment x *Kharif* rainfall x loam soil	-	0.00425
		(0.00758)
H_0: Investment effect = 0, $\chi^2(n)$ [p]	11.5(2)	14.3(4)
	[.0031]	[.0063]
Investment effect at sample mean investment and rainfall	6.41	6.94
	(2.12)	(2.84)
N (farmer-years)	1,152	1,152

Standard errors clustered at the farm level in parentheses.
Specification includes rainfall and lagged rainfall and rainfall and lagged rainfall interacted with total landholdings, irrigated landholdings, and four soil characteristics; and village*year fixed effects.
The identifying instruments for the planting-stage investment variables are the year-specific IMD Southern Peninsula pre-planting monsoon forecast interacted with total landholdings, irrigated landholdings, and four soil characteristics.

Table 2
Probability of Being within ± δ of the Expected Return on Investment, by δ and Sample

δ (Percentage Points)	ICRISAT	Ghana
10	.012	.02
20	.02	.03
30	.04	.04
40	.05	.05
50	.06	.06

The probabilities are calculated based on equations (3) and (4) in the text. In column 1 we use the estimate of α (the interaction between investment and the rainfall realization, evaluated at mean investment) from Table 1, column 2, multiplied by the estimated standard deviation of the rainfall distribution in the ICRISAT village with the longest rainfall series. In column 2, we use the estimates of α from column 2 of Table 3, evaluated at the mean investment level, and the empirical distribution of the bivariate weather outcomes in the historical rainfall distribution of the Ghana study area.

Table 3
Farmer Fixed Effects IV Estimates of Planting-Stage Investments on Net Returns,
75 Ghana Villages, 2009-2011

Sample years	2009-2011		2009-2010
Variable	(1)	(2)	(3)
Planting-stage investment	1.043 (0.707)	0.956 (0.718)	-0.0144 (0.766)
Planting-stage investment x rainfall index	-	0.0298 (0.0119)	0.0344 (0.0123)
Planting-stage investment x total rainfall	-	0.123 (0.0586)	0.128 (0.0569)
H_0: Investment effect = 0, $\chi^2(3)$ [p]	-	11.3 [.0104]	7.84 [.0494]
H_0: Rainfall-investment interactions = 0, $\chi^2(2)$ [p]	-	6.32 [.0424]	7.78 [.0204]
N (farmer-plot-years)	6,418	6,418	4,126

Standard errors clustered at the farmer level in parentheses. The specifications include the rainfall index and its square, five soil characteristics, five slope characteristics, these characteristics interacted with the rainfall index, the distance of the plot to the house, the number of trees, total rainfall, total rainfall and the rainfall index interacted., plot area, and year. The identifying instruments for the planting-stage variables are randomized offers of index insurance at nine price levels, cash grants, and cash grants plus insurance.

Table 4
Cross-Sectional IV Estimates of Planting-Stage Investments on Net Returns, by Year
75 Ghana Villages, 2009-2011

Variable/Sample	2009	2010	2011
Planting-stage investment	1.402	0.720	1.715
	(6.27)	(0.950)	(1.80)
Planting-stage investment x rainfall index	0.0391	-0.00121	0.0702
	(0.0645	(0.0118)	(0.0296)
Planting-stage investment x total rainfall	0.176	0.0544	0.169
	(0.0858)	(0.0826)	(0.127)
H_0: Investment effect = 0, $\chi^2(3)$ [p]	6.28	1.54	11.4
	[.0990]	[.6737]	[.0097]
H_0: Rainfall-investment interactions = 0, $\chi^2(2)$ [p]	4.62	0.450	5.64
	[.0993]	[.7979]	[.0597]
N (farmer-plot-years)	1,844	2,282	2,292

Standard errors clustered at the farmer level in parentheses. The specifications include the rainfall index and its square, five soil characteristics, five slope characteristics, these characteristics interacted with the rainfall index, the distance of the plot to the house, the number of trees, total rainfall, total rainfall and the rainfall index interacted., plot area, and year. The identifying instruments for the planting-stage variables are randomized offers of index insurance at nine price levels, cash grants, and cash grants plus insurance.

Table 5
Enterprise Fixed Effects IV Estimates of Investments on Net Returns, by Survey Wave
Sri Lankan Micro-enterprises, 2004-2006

Variable	(1)	(2)
Investment (non land)	0.0756 (0.0242)	-
Entrepreneur sick	-1042.8 (329.4)	-1033.1 (391.9)
Wave 2 investment	-	0.133 (0.0718)
Wave 3 investment	-	0.132 0(.0503)
Wave 4 investment	-	0.0615 (0.0315)
Wave 5 investment	-	0.0795 (0.0311)
Wave 6 investment	-	0.0664 (0.0252)
Wave 7 investment	-	0.0794 (0.0267)
Wave 8 investment	-	0.0701 (0.0289)
Wave 9 investment	-	0.0884 (0.0267)
Wave fixed effects	Y	Y
H_0: wave fixed effects=0, $\chi^2(8)$ [p]	30.1 [.0002]	5.45 [.709]
H_0: Seasonal pattern, $\chi^2(4)$ [p]	-	2.22 [.694]
N (enterprise-quarters)	3,101	3,101

Standard errors clustered at the enterprise level in parentheses.
Sample excludes enterprises directly affected by the 2004 tsunami.
The identifying instruments in column 1 are the randomized offer of cash and that variable interacted with whether the entrepreneur was a male. In column 2 the instruments also include the two instruments interacted by wave indicators.

Table 6
Enterprise Fixed Effects IV Estimates of Investments on Net Returns, by Quarterly Season
Sri Lankan Micro-enterprises, 2004-2006

Variable	(1)	(2)
Investment (non land)	-	0.0977
		(0.0216)
Entrepreneur sick	-1066.5	-951.5
	(366.5)	(345.3)
Q1 investment	0.0947	-
	(0.0219)	
Q2 investment	0.1001	-
	(0.0236)	
Q3 investment	0.0693	-
	(0.0192)	
Q4 investment	0.0836	-
	(0.0242)	
Season fixed effects	Y	Y
H_0: Season fixed effects = 0 $\chi^2(4)$ [p]	9.27 [.0547]	11.6 [.0208]
N (enterprise-quarters)	3,101	3,101

Standard errors clustered at the enterprise level in parentheses.
Sample excludes enterprises directly affected by the 2004 tsunami.
The identifying instruments in the first column are the randomized offer of cash and that variable interacted with whether the entrepreneur was a male and those two instruments interacted with wave indicators. The identifying instruments in column 2 are the randomized offer of cash and that variable interacted with whether the entrepreneur was a male and quarter indicators.

Table 7
Log Urban Hourly Wages: Estimates of the Returns to Schooling by Year, 1993-2000
Men Aged 16-79 in Urban Indonesia (IFLS 1993, 1997, 2000)

Variable/Estimation Procedure	OLS	OLS	FE
Years of schooling	0.123 (0.00485)	-	-
Years of schooling/ Difference from 2000, 1993	-	0.154 (.0103)	-0.00360 (0.0160)
Years of schooling/ Difference from 2000, 1994	-	0.124 (0.00730)	-0.0129 (0.1013)
Years of schooling/ Difference from 2000, 1995	-	0.121 (0.00692)	-0.0139 (0.0102)
Years of schooling/ Difference from 2000, 1996	-	0.122 (0.00677)	-0.0125 (0.0102)
Years of schooling/ Difference from 2000, 1997	-	0.138 (0.00768)	-0.000132 (0.0104)
Years of schooling/ Difference from 2000, 1998	-	0.104 (0.00651)	-0.0255 (0.00752)
Years of schooling/ Difference from 2000, 1999	-	0.102 (0.00623)	-0.0270 (0.00707)
Years of schooling 2000	-	0.129 (0.00761)	-
Year fixed effects	Y	Y	Y
H_0: Year fixed effects=0, F(8, 3062) [p]	415.0 [.0000]	395.1 [.0000]	17.0 [.0000]
H_0: Schooling returns are equal across years, F(7,3062) [p]	-	5.94 [.0000]	2.95 [.0044]
N (person-years)	12,769	12,769	12,769

Standard errors clustered at the individual level in parentheses.
Specifications also include age and age squared.

Table 8
Log Urban Hourly Wages: Estimates of the Returns to Schooling and External Macro Shocks
Men Aged 16-79 in Urban Indonesia (IFLS 1993, 1997, 2000)

Sample Period	1993-2000		1998-2000	
Estimation Procedure	OLS	FE	OLS	FE
Years of schooling	0.0855 (0.0102)	-	0.00134 (.0213)	-
Years of schooling x oil price	0.00105 (0.000492)	0.00156 (0.000434)	0.00214 (0.000517)	0.00210 (0.000631)
Years of schooling x exch. rate x industry openness x ind. skill x 10^{-5}	-	-	0.222 (0.0506)	0.121 (0.0615)
Oil price	-.0528 (.0136)	-0.0504 (0.0106)	-0.0894 (0.0131)	-0.0632 (0.0196)
Exchange rate x industry openness x skill x10^{-4}	-	-	-0.342 (0.107)	-0.613 (0.258)
Exchange rate x10^{-4}	-	-	0.447 (0.209)	0.437 (0.746)
Industry openness x skill	-	-	0.0968 (0.0533)	-
H_0: No shock effects on schooling returns, F(2, 2120) [p]	-	-	13.3 [.0001]	5.54 [.0040]
N (person-years)	12,242	12,242	5,470	5,470

Standard errors clustered at the individual level in parentheses.
Specifications also include age and age squared, year, and age and age squared interacted with the oil price/industry variables.

Appendix Table A1
Estimates of the Effects of Year-Specific Rainfall and Mean Rainfall on Log Total Household Per-capita Wealth at Baseline (2009)

Variable/rainfall years	2000-2015	2009	2010	2011
Rainfall in year	-	-0.731	11.8	0.0193
		(0.906)	(4.35)	(0.00565)
Rainfall in year squared	-	0.0568	-0.767	-0.0000117
		(0.0595)	(0.285)	(0.0000037)
Mean rainfall	0.263	-	-	-
	(0.0193)			
Mean rainfall squared	-0.0000152	-	-	-
	(0.0000122)			
Number of farmers	1,065	1,065	1,065	1,065
H_0: No rainfall effect F(2,68)	4.70	4.62	3.89	10.37
[p]	[0.0123]	[0.0132]	[0.0251]	[0.0001]

Standard errors clustered at the village level in parentheses.

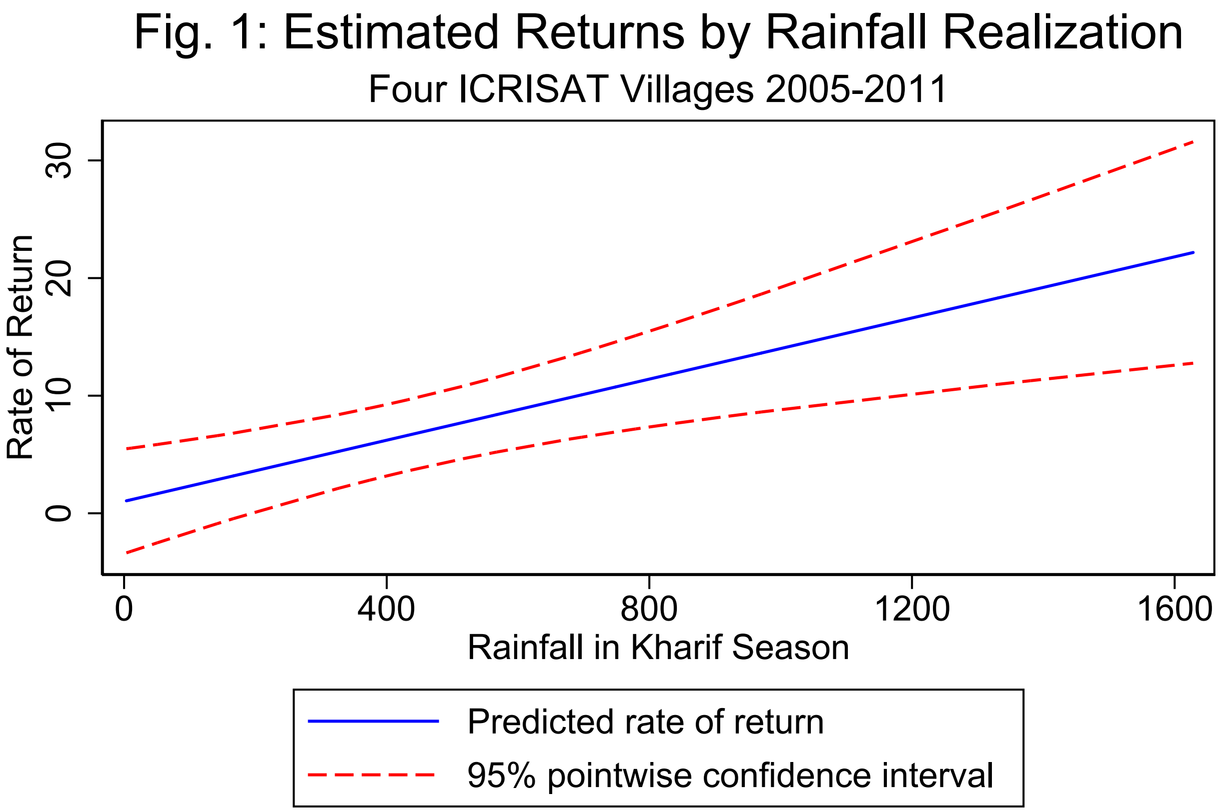


Fig. 1: Estimated Returns by Rainfall Realization
Four ICRISAT Villages 2005-2011
Rate of Return
0
10
20
30
0
400
800
1200
1600
Rainfall in Kharif Season
Predicted rate of return
95% pointwise confidence interval

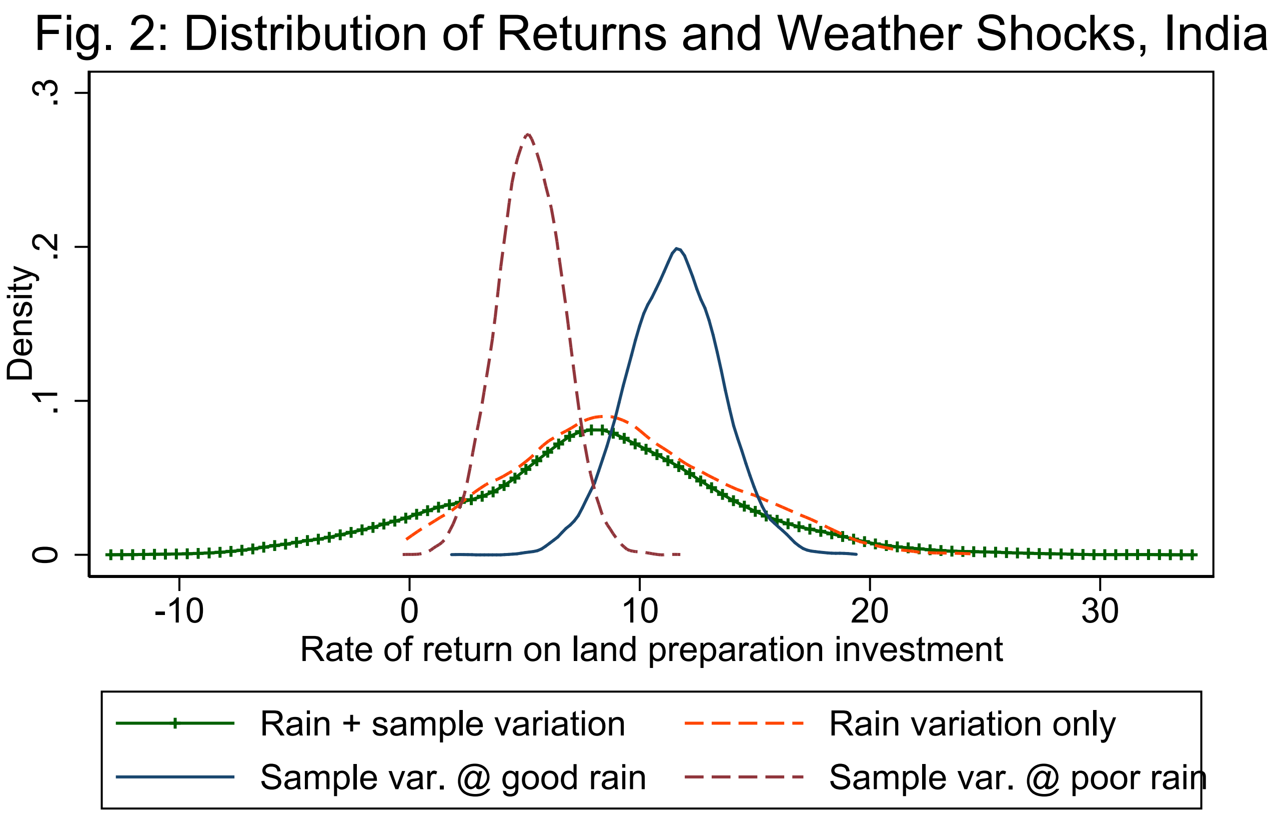


Fig. 2: Distribution of Returns and Weather Shocks, India
Density
0
.1
.2
.3
-10
0
10
20
30
Rate of return on land preparation investment
Rain + sample variation
Rain variation only
Sample var. @ good rain
Sample var. @ poor rain

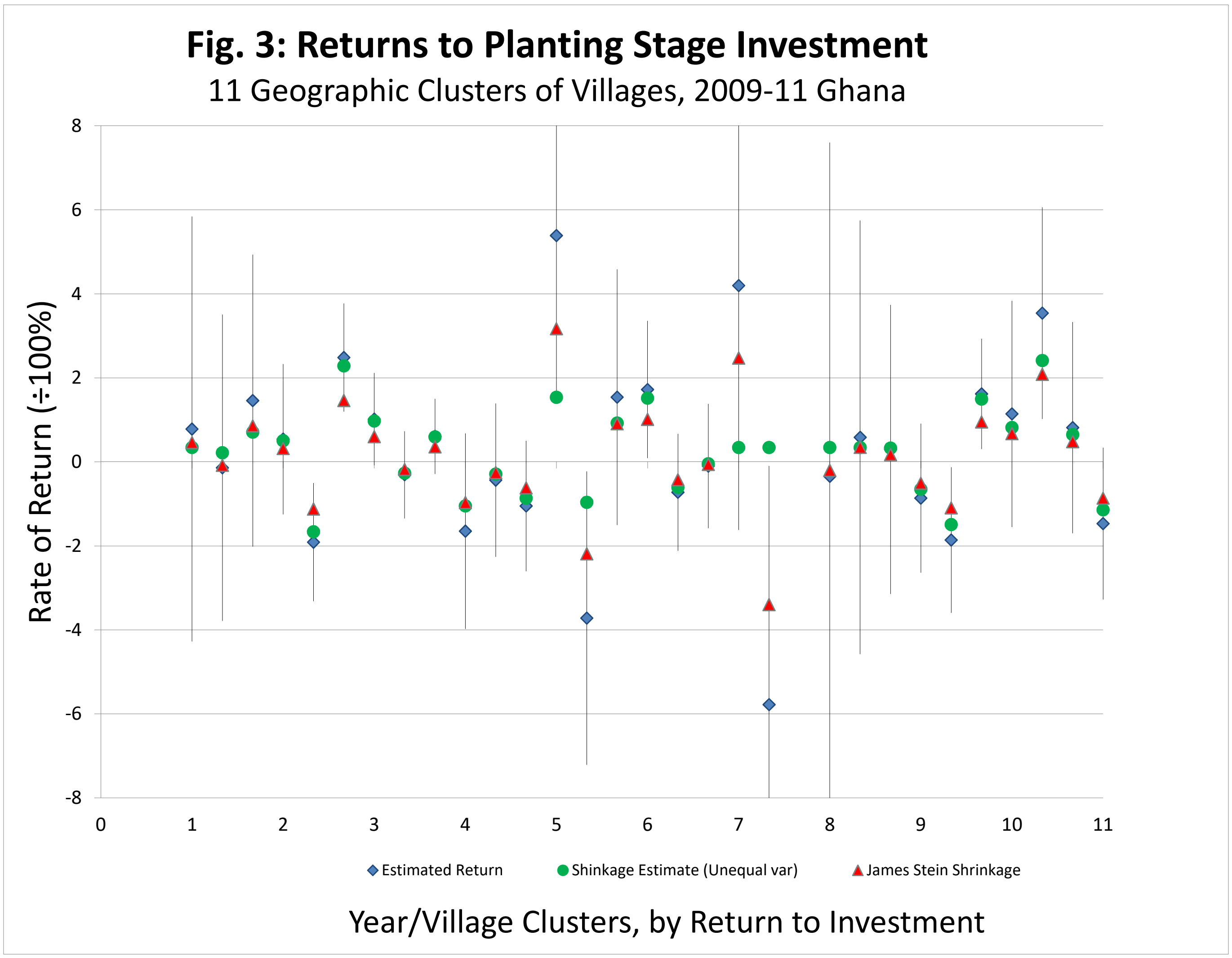
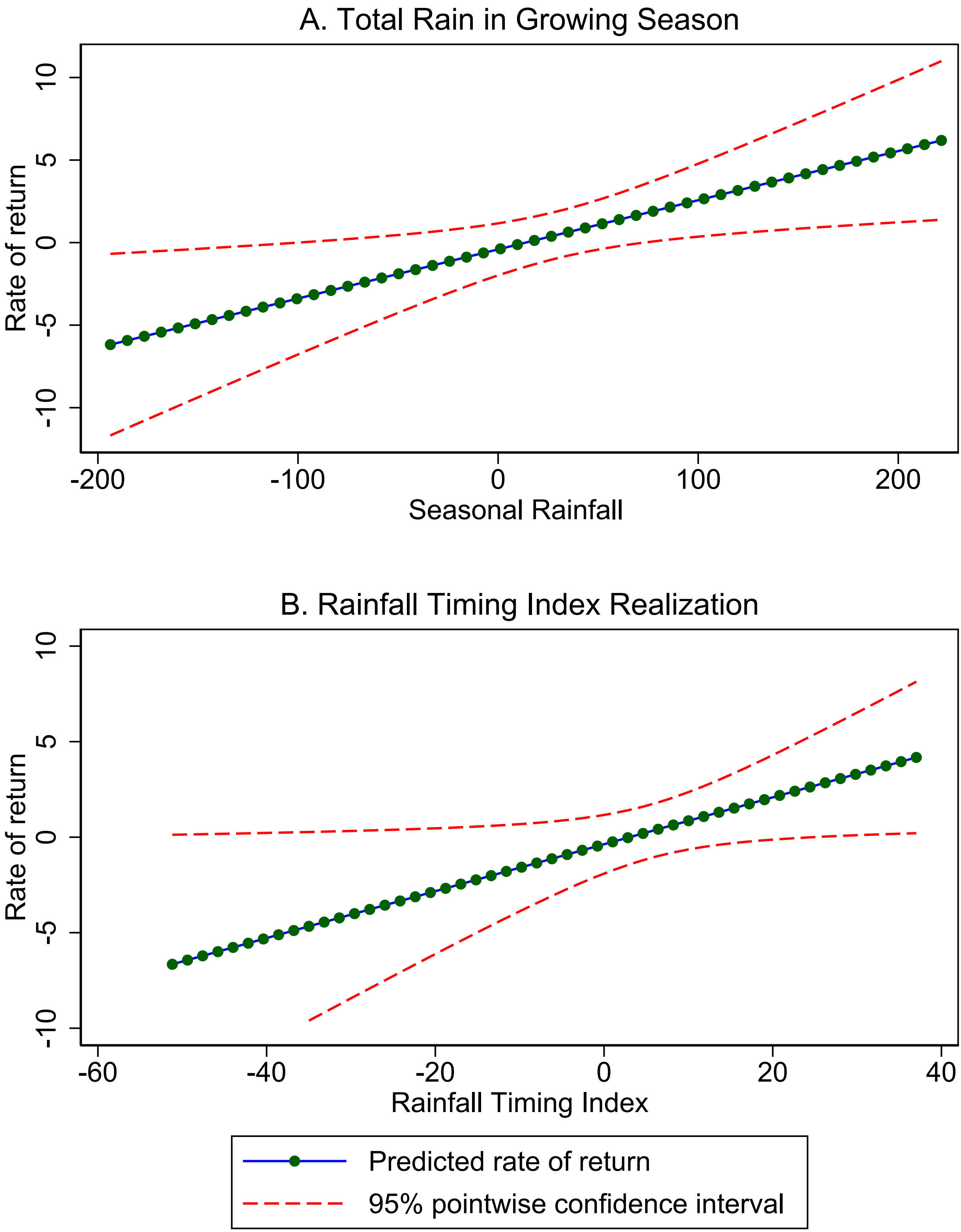


Fig. 3: Returns to Planting Stage Investment
11 Geographic Clusters of Villages, 2009-11 Ghana
Rate of Return (÷100%)
8
6
4
2
0
-2
-4
-6
-8
0
1
2
3
4
5
6
7
8
9
10
11
Estimated Return
Shinkage Estimate (Unequal var)
James Stein Shrinkage
Year/Village Clusters, by Return to Investment

Fig. 4: Estimated Returns by Rainfall Realization



Across 75 Northern Ghana Villages, 2009-2011. Std. errors clustered at community level.
Evaluated at mean seasonal rainfall.

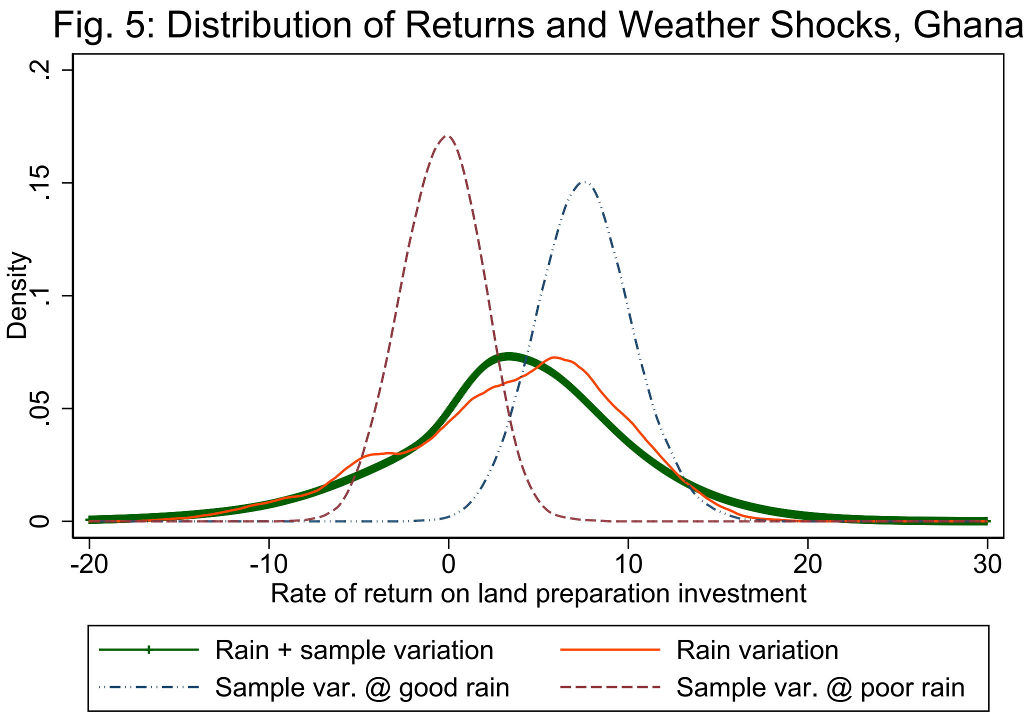


Fig. 5: Distribution of Returns and Weather Shocks, Ghana
Density
0
.05
.1
.15
.2
-20
-10
0
10
20
30
Rate of return on land preparation investment
Rain + sample variation
Rain variation
Sample var. @ good rain
Sample var. @ poor rain

Fig. 6: Estimated Returns by Rainfall Realization

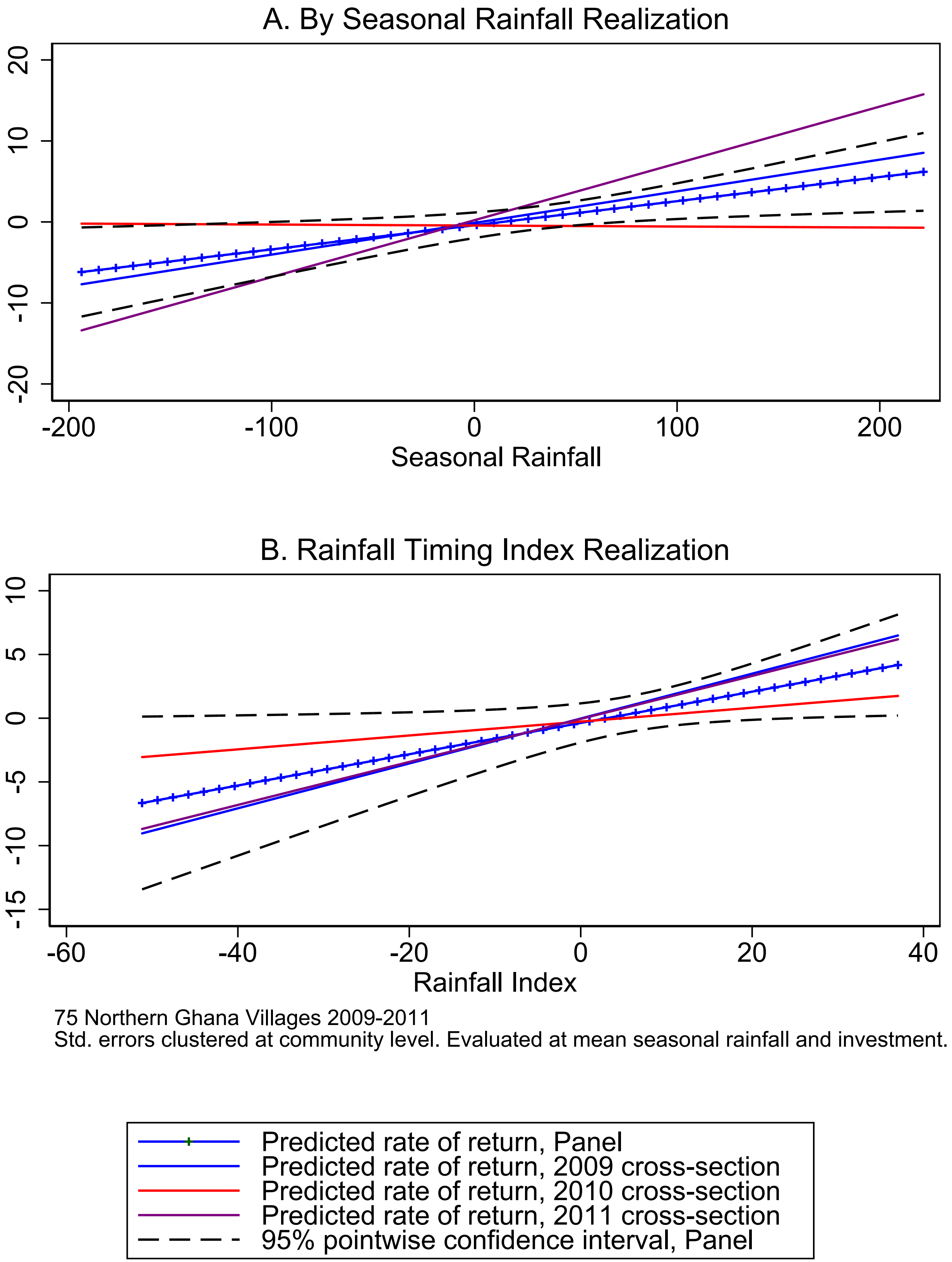


Figure 7. Oil Price and Exchange Rate Fluctuations, Indonesia 1990-2000

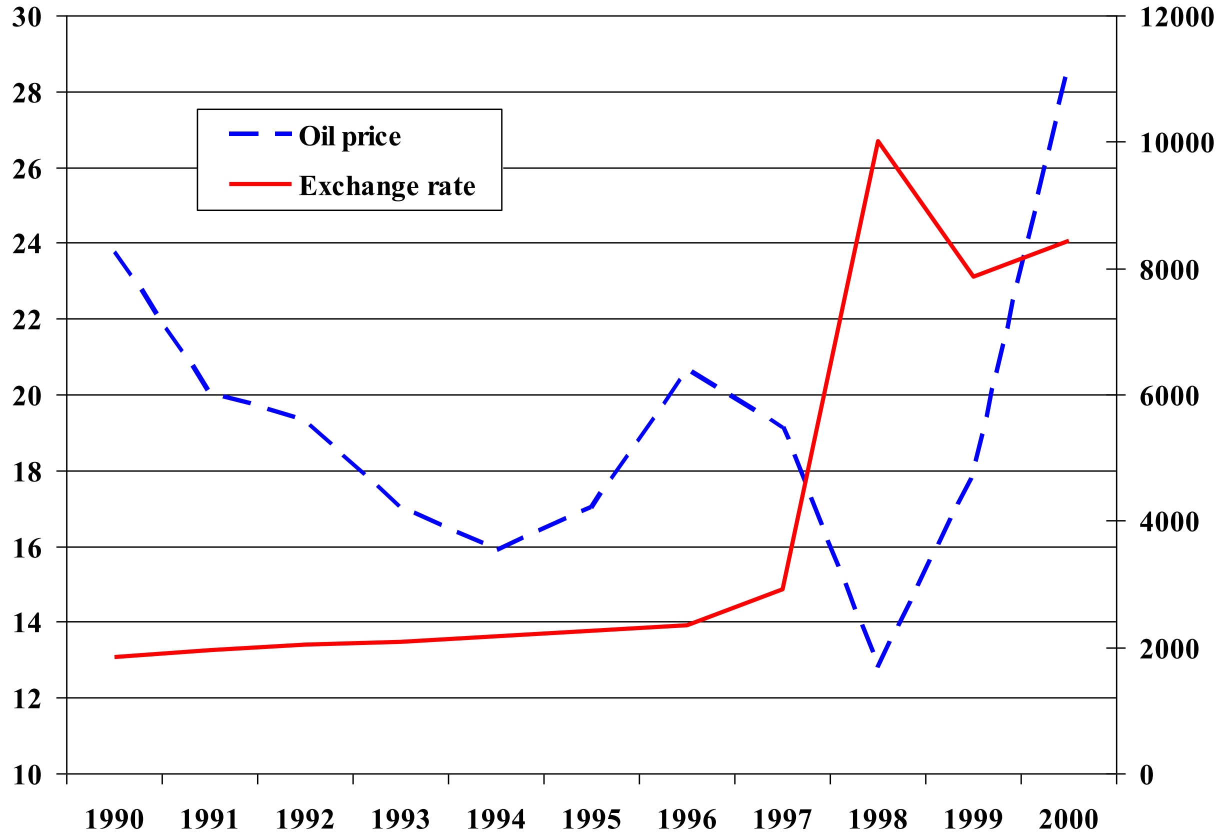


Figure 8. Average Annual OPEC crude oil price: 1960 - 2016(U.S. dollars per barrel)

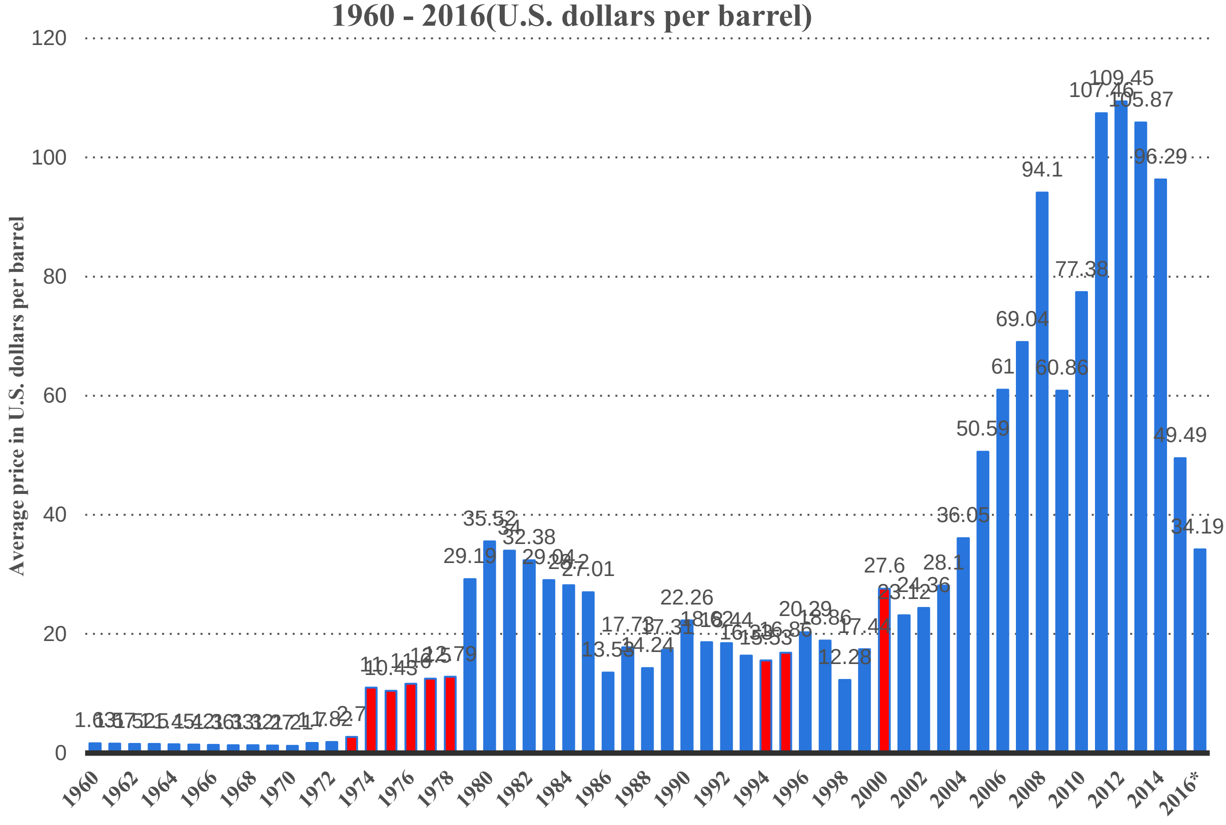
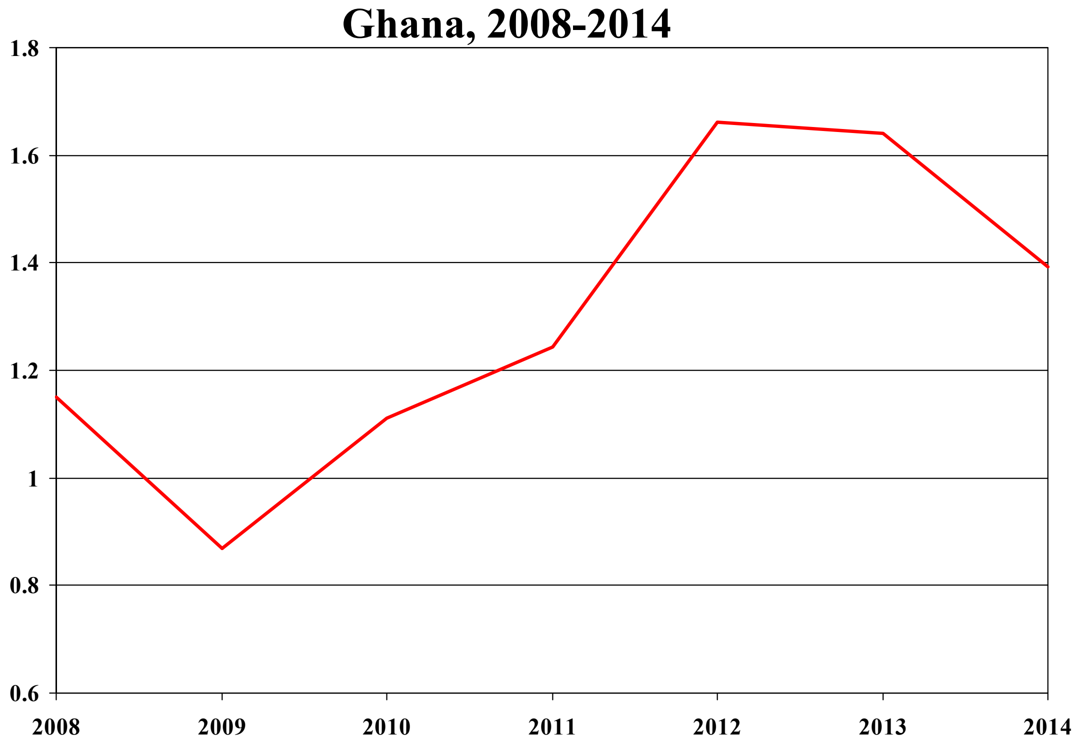
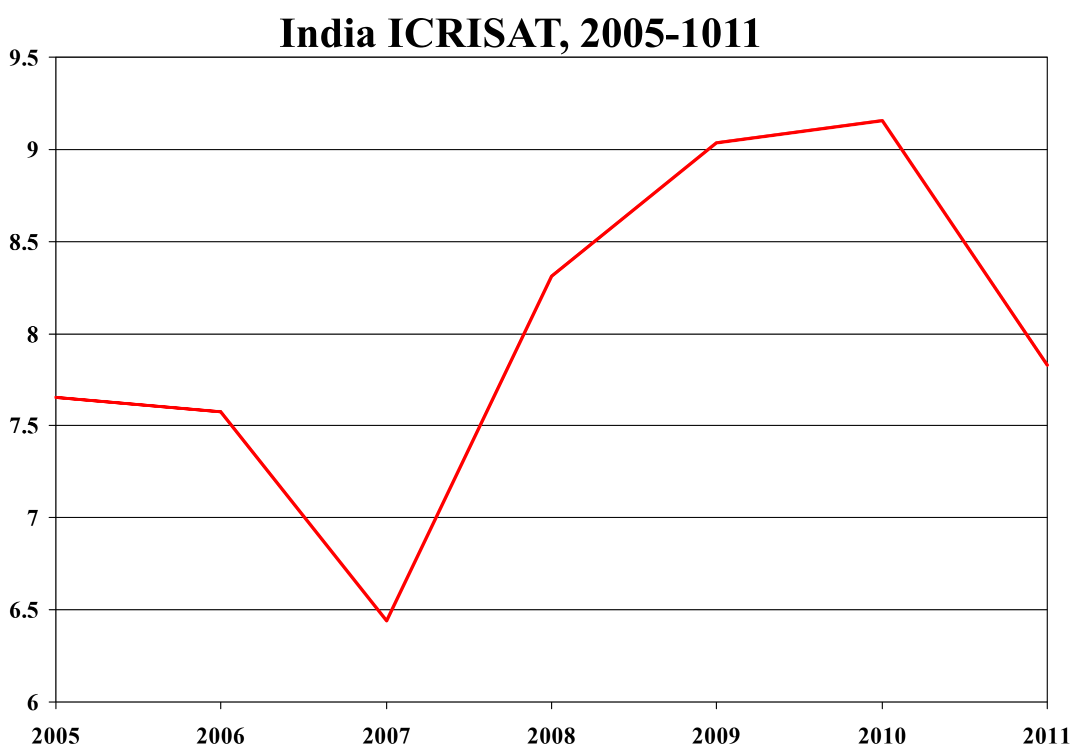


Figure A1. Investment Return Variation Due to Output Price Variation





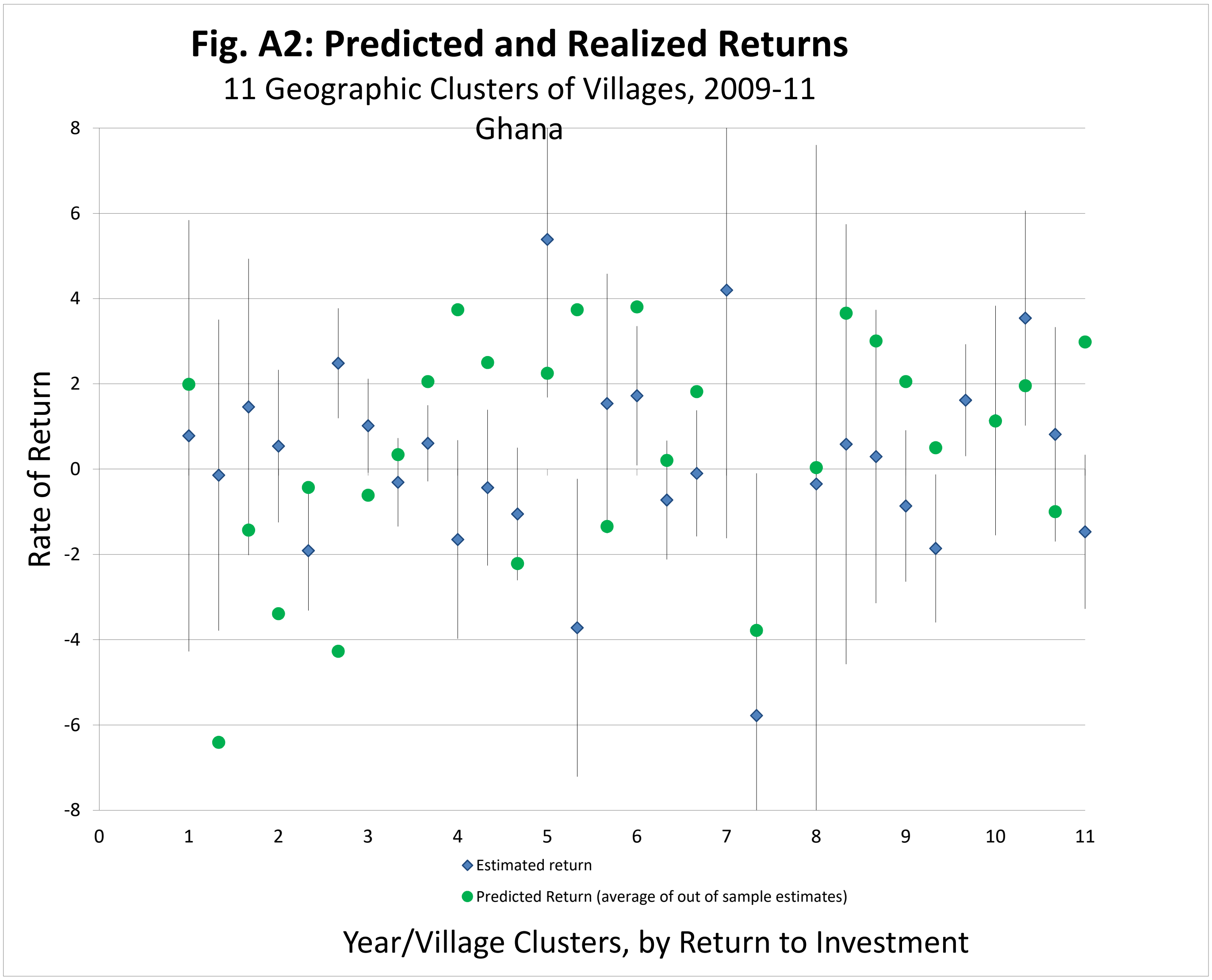


Fig. A2: Predicted and Realized Returns
11 Geographic Clusters of Villages, 2009-11
Ghana
Rate of Return
8
6
4
2
0
-2
-4
-6
-8
0
1
2
3
4
5
6
7
8
9
10
11
Estimated return
Predicted Return (average of out of sample estimates)
Year/Village Clusters, by Return to Investment